

**REMARKS/ARGUMENTS**

Claims 18-30 are currently pending in the above-identified application. No claims have been amended in the present response. Applicants respectfully request reconsideration of the pending claims in light of the remarks below.

Rejections under 35 U.S.C. § 102:

Claims 18, 19 and 22 remain rejected under 35 U.S.C. § 102(b) as being anticipated by U.S. Patent 5,338,686 (Hellerstein, "the '686 patent"). The Examiner asserts that Hellerstein teaches a method for determining the rate of degradation of an isotopically labeled biopolymer comprising adding a stable isotope-labeled subunit (monomer) to a biopolymer pool, collecting first and second samples and measuring the relative abundance of monoisotopic and isotopomeric peaks using mass spectrometry, calculating the difference between the peaks of the first and second samples and determining the rate of biopolymer decay. The Examiner further alleges that serial time points were collected and individual mass isotopomers were plotted over time to determine the rate of decay. The Examiner concludes from these alleged actions that Hellerstein teaches every element of the claimed invention.

In addition, the Examiner questions Applicants assertion that the calculation of molar excesses is not the same as the determination of abundance. The Examiner alleges that Hellerstein calculates relative abundance using mass spectrometry and further remarks that she is unclear what the difference is between molar excess and relative abundance, since both are looking at abundance of monoisotopic and isotopomeric peaks. Further, the Examiner returns to the Hellerstein's statement that mass spectral peak heights may be expressed as ratios for purposes of determining relative frequencies, but also states that any calculation means which provides relative values for abundance of such isotopomers in a sample may be used in describing the data.

Applicants must again respectfully disagree with the Examiner rejection. The Examiner has questioned the difference between molar excess and relative abundance and the relevance of the difference to any patentable differences between Hellerstein and the present invention. Molar excess as described and used by Hellerstein describes a calculation of the observed distribution of a mass isotopomer in a polymer containing at least two identical subunits. In calculating molar excess Hellerstein first isolated each biopolymer containing 1, 2, 3, 4, . . . etc, ( $M_0$ ,  $M_1$ ,  $M_2$ , . . . etc.) isotopically labeled subunits. Each of the isolated calculated values is for a particular isotopomer species. The determined value for each isotopomer species is then compared to a probability matrix that is created on assumptions of precursor pool percentages, which in the method of Hellerstein results in the calculation of the molar excess for each isotopomer species. Once the molar excess is calculated, the frequency of replacement for each isotopomeric species is compared at two different time points to calculate the decay of a biopolymer. In a 1999 review published by the inventor, (Hellerstein and Neese, *Am. J. Physiol.* 276:E1146-1170, 1999), state "that subpopulations of molecules must be distinguishable and quantifiable" (page E1147, left column, fourth full paragraph, lines 1 and 2). In addition, they state that "there must be combinations possible in the molecule analyzed" (page E1147, left column, third full paragraph, lines 1 and 2) and that "the polymers must be intact, or with at least two subunits present, because the distribution of the isotopomeric species carries the essential information." (page E1147, left column, third full paragraph, lines 7 - 10). This characteristic of the invention of Hellerstein is recited in column 6, lines 60 through 64.

In order to try to further clarify the method of Hellerstein, Applicants provide the following analogy also used by Hellerstein in publications from his laboratory. The analogy uses a necklace containing colored beads to represent a biopolymer comprising a series of subunits. In the analogy, specific beads (within a pattern of beads of many colors that make up the whole necklace) can either be colored white (non-labeled) or colored black (isotopically-labeled). Naturally occurring necklaces (biopolymers) will contain white beads and some percentage of black beads where white beads should be found because black beads occur naturally (naturally-occurring stable isotopes found in nature). When the bead precursor pool is enriched with black

beads over white beads, there is an increased probability that black beads will be incorporated into the necklaces at the location where white beads are usually found. The method of Hellerstein looks at individual necklaces, isotopomeric species, using, in the specific examples gas chromatography, measuring the number of necklaces containing 1, 2, 3 ... *etc.* black beads, and comparing these data to probability matrices calculated using different proportions of white and black beads available for stringing on necklaces (the precursor pool). Further, Hellerstein uses the values measured for the mass isotopomeric peaks to determine whether the biopolymers are substituted with 1, 2, 3, ..., *etc.* isotopically-labeled subunits.

Fig. 1B of the '686 patent provides a description of the Hellerstein method for measuring isotopic decay. In the first box, the molar excess for each necklace, isotopomeric species, the polymer having one substitution  $M_1$  through the polymer having four substitutions  $M_4$  is calculated at each time point. In the second step the ratio (R) of the calculated molar excess to the sum of the molar excesses for each necklace is calculated. Time points and masses for which "R" is stable are used to calculate the decay constant (k). The statement of Hellerstein does not disclose or suggest the method of the present invention.

In the method of the present invention, each isotopomer species is not analyzed separately, each necklace, but instead the pooled distribution of isotopomers, or necklaces or necklace fragments, is analyzed. There is no need to distinguish subpopulations of polymers as data is collected on a pooled population containing both labeled and unlabeled polymer (relative abundance). This simpler method can be used even when only one subunit is present within the polymer being analyzed. In one aspect of the method of the present invention a necklace (biopolymer) with a known pattern of white beads is broken down into fragments. Some of the fragments will contain white beads that can be substituted with black beads (A fragments), and other fragments will not contain any white beads (B fragments). One or more of the necklace fragments must contain white beads (or black beads) for our method to work, but unlike Hellerstein there is no need for two repeats of an identical subunit to be present in the biopolymer. In the method of the present invention, a pool of specific A fragments is analyzed, and the mass isotopomeric distribution of the pooled population is measured by mass

spectrometry. In order to calculate biopolymer formation or degradation, the mass isotopomeric distribution of A fragments that have been formed in the presence of large excesses of black beads is compared to the mass isotopomeric distribution of pooled control A fragments (containing a mixture of white beads and naturally-occurring black beads). The percentage of newly formed necklaces is determined by mathematical comparisons of mass isotopomeric abundances of pooled necklace fragments before and after enrichment with black beads.

As such, Hellerstein does not teach every element of the claimed invention. As above, Hellerstein instead teaches a method that calculates the molar excess of each isotopomer species and the frequency for that species. This value is compared at different time points. The calculation of the molar excess and frequency for each isotopomer does not disclose or suggest the present invention.

The Examiner has also again referred to a comment in Hellerstein that any calculation means which provides relative values for abundance of such isotopomers in a sample may be used in describing the data as teach elements of the present invention. Applicants respectfully must again point out to the Examiner that this statement is made in the context of calculating the molar excess for each isotopomer species and not the pooled relative abundance for all monoisotopic and isotopomeric peaks for an intact or fragment of a biopolymer. The present invention provides a simpler method for determining a degradation rate for a biopolymer that is not disclosed or suggested by Hellerstein.

Applicants note that claims 20, 21 and 23 through 30 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten in independent form including all of the limitations of the base claim and any intervening claims.

In view of the above remarks, the Examiner is respectfully requested to reconsider and withdraw the rejection of claims 18, 19 and 22 under 35 U.S.C. § 102(b) as being anticipated by U.S. Patent 5,338,686 (Hellerstein).

B. ROBERT FRANZA JR. *et al.*  
Application No. 10/722,161  
Examining Group 1636

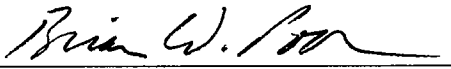
PATENT

**CONCLUSION**

In view of the foregoing, Applicants believe all claims now pending in this Application are in condition for allowance and an action to that end is respectfully requested. If the Examiner believes a telephone conference would expedite prosecution of this application, please telephone the undersigned at 206-467-9600.

Respectfully submitted,

Dated: 27 December 2007

  
\_\_\_\_\_  
Brian W. Poor  
Reg. No. 32,928

TOWNSEND and TOWNSEND and CREW LLP  
Two Embarcadero Center, Eighth Floor  
San Francisco, California 94111-3834  
Tel: 206-467-9600  
Fax: 415-576-0300  
BWP/jlv  
61228647 v1

# Mass isotopomer distribution analysis at eight years: theoretical, analytic, and experimental considerations

MARC K. HELLERSTEIN<sup>1,2</sup> AND RICHARD A. NEESE<sup>1,2</sup>

<sup>1</sup>Department of Nutritional Sciences, University of California at Berkeley, Berkeley 94720;

and <sup>2</sup>Division of Endocrinology and Metabolism, Department of Medicine, University of California, San Francisco, California 94110

**Hellerstein, Marc K., and Richard A. Neese.** Mass isotopomer distribution analysis at eight years: theoretical, analytic, and experimental considerations. *Am. J. Physiol.* 276 (*Endocrinol. Metab.* 39): E1146–E1170, 1999.—Mass isotopomer distribution analysis (MIDA) is a technique for measuring the synthesis of biological polymers. First developed approximately eight years ago, MIDA has been used for measuring the synthesis of lipids, carbohydrates, and proteins. The technique involves quantifying by mass spectrometry the relative abundances of molecular species of a polymer differing only in mass (mass isotopomers), after introduction of a stable isotope-labeled precursor. The mass isotopomer pattern, or distribution, is analyzed according to a combinatorial probability model by comparing measured abundances to theoretical distributions predicted from the binomial or multinomial expansion. For combinatorial probabilities to be applicable, a labeled precursor must therefore combine with itself in the form of two or more repeating subunits. MIDA allows dilution in the monomeric (precursor) and polymeric (product) pools to be determined. Kinetic parameters can then be calculated (e.g., replacement rate of the polymer, fractional contribution from the endogenous biosynthetic pathway, absolute rate of biosynthesis). Several issues remain unresolved, however. We consider here the impact of various deviations from the simple combinatorial probability model of biosynthesis and describe the analytic requirements for successful use of MIDA. A formal mathematical algorithm is presented for generating tables and equations (APPENDIX), on the basis of which effects of various confounding factors are simulated. These include variations in natural isotope abundances, isotopic disequilibrium in the precursor pool, more than one biosynthetic precursor pool, incorrect values for number of subunits present, and concurrent measurement of turnover from exogenously labeled polymers. We describe a strategy for testing whether isotopic inhomogeneity (e.g., an isotopic gradient or separate biosynthetic sites) is present in the precursor pool by comparing higher-mass (multiply labeled) to lower-mass (single- and double-labeled) isotopomer patterns. Also, an algebraic correction is presented for calculating fractional synthesis when an incomplete ion spectrum is monitored, and an approach for assessing the sensitivity of biosynthetic parameters to measurement error is described. The different calculation algorithms published for MIDA are compared; all share a common model, use overlapping solutions to computational problems, and generate identical results. Finally, we discuss the major practical issue for using MIDA at present: quantitative inaccuracy of instruments. The nature and causes of analytic inaccuracy, strategies for evaluating instrument performance, and guidelines for optimizing accuracy and reducing impact on biosynthetic parameters are suggested. Adherence

to certain analytic guidelines, particularly attention to concentration effects on mass isotopomer ratios and maximizing enrichments in the isotopomers of interest, reduces error. Improving instrument accuracy for quantification of isotopomer ratios is perhaps the highest priority for this field. In conclusion, MIDA remains the “equation for biosynthesis,” but attention to potentially confounding factors and analytic performance is required for optimal application.

THE ASSEMBLY AND DISASSEMBLY of polymers synthesized from repeating monomeric units is a central theme in biology. Such polymers may be as simple as fatty acids synthesized from acetyl-CoA units or as complex as proteins synthesized from amino acids or DNA made from nucleotides. Other examples include carbohydrates (e.g., glucose from triose units, glycogen from glucose, glycoproteins), porphyrins (e.g., chlorophyll, heme), and lipids (e.g., cholesterol, triacylglycerols). Biological polymers may be homonuclear (defined as containing subunits that are identical), as in fatty acids, or heteronuclear (defined as containing more than one type of subunit), as in proteins or polynucleotides. Despite the importance of polymers in the chemistry of living systems, techniques for determining their rates of synthesis or breakdown have historically been unsatisfactory (1, 9, 18, 19). As a consequence, fields as wide ranging as lipid biosynthesis, protein metabolism, carbohydrate metabolic regulation, and control of cell proliferation have been severely constrained.

In this article, we will provide an update and eight-year perspective on a technique that provides a fundamental solution to the problem of measuring polymerization biosynthesis. Mass isotopomer distribution analysis (MIDA) is a technique based on combinatorial probabilities and the labeling patterns in intact polymers that can be said to provide a fundamental “equation for biosynthesis.” Although MIDA was first presented as a systematic approach to polymerization biosynthesis only a few years ago (13–15, 20), a number of refinements, alternative calculation algorithms, and criticisms have been published since then (4, 5, 22, 31, 32). We will review here the theoretical and practical factors that must be taken into account if MIDA and related techniques are to be applied successfully.

## BIOLOGICAL BASIS OF MEASUREMENT OF POLYMERIZATION BIOSYNTHESIS BY ISOTOPE INCORPORATION

The principle of isotope incorporation techniques for measuring polymerization biosynthesis is, on the sur-

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

face, straightforward. In a biological system, polymers that are newly synthesized mix into a pool that also contains preexisting polymer molecules. The goal of an isotope incorporation study is to quantify the fraction of molecules in the mixture that were newly synthesized during the label incorporation period (i.e., "what's new") and the rate at which the total pool of polymers is turning over. To determine the newly synthesized fraction ( $f$ ) present in the mixture, one must first establish exactly how much label is contained in the population of newly synthesized polymers. Dilution of this labeled population by the population of preexisting, unlabeled molecules can then be determined, according to the precursor-product relationship (15, 22, 39, 40).

The major practical difficulty has been establishing how much label is contained in the newly synthesized population of molecules. There exists no purely physical technique for identifying in a mixed population of molecules which ones are new and which are not. No classical extraction technique can reveal where different molecules in a population came from or how long they have been present. The biochemistry of the precursor-product relationship provides a possible solution, however (Fig. 1), because the precursor pool of subunits in a cell has a physical reality and can in principle be isolated by extraction techniques.

Serious problems arise when investigators have tried to use surrogate monomer pools to represent the isotopic content of the true precursor pool ( $p$ ) (9, 19, 34, 37, 38), however: e.g., plasma amino acids or free intracellular amino acids to represent the tRNA-amino acid precursor pool for protein synthesis, or ketone bodies to represent the acetyl-CoA pool for lipogenesis. Complicating factors deriving from subcellular or intracellular biochemical organization have been shown to affect every class of polymer so far examined in detail, including proteins (38), lipids (8, 9), carbohydrates (19, 34), and nucleic acids (18).

#### A SOLUTION TO THE PRECURSOR-PRODUCT PROBLEM: THE USE OF COMBINATORIAL PROBABILITIES

MIDA is based on a model of combinatorial probabilities. Polymerization biosynthesis can be conceptualized as a combinatorial process, with monomeric subunits from a precursor pool combining into a polymeric collection or assemblage. If the monomeric subunits are of more than one distinctive type, i.e., labeled and unlabeled, then the population of assembled polymers will not be of uniform isotopic composition. The polymers will exist as distinguishable species containing varying numbers of the different types of subunits. Some species will include no labeled subunits, some will include one labeled subunit, some will contain two labeled subunits, and so on. The relative proportion of each species of polymer is determined by and can be calculated from the binomial (or multinomial) expansion (Fig. 1A). The binomial expansion contains two variables, the number of subunits in the collection ( $n$ ) and the probability ( $p$ ) of each subunit being of a particular type. Because the number of subunits in a biological polymer is constant and known, the sole

factor determining the relative proportions of each polymeric combination (i.e., the quantitative distribution of mass isotopomers) is  $p$ , the labeling probability in the precursor pool.

The population of intact polymeric assemblages therefore contains information about the precursor pool that is not available by analysis of the monomeric units in isolation: the combinations of labeled and unlabeled subunits in the polymer population, which are manifested for statistical or mathematical analysis as the distribution of mass isotopomers. This is the central insight on which MIDA is based. Because each isotopomeric distribution is uniquely determined by  $p$ , each distribution is characteristic of and capable of revealing the unique value of  $p$  from which it was assembled. The distribution is, moreover, immutable; it is a fingerprint that will persist throughout the lifetime of the population, as long as there is no biological discrimination (no isotope effect) between species of the polymer and no remodeling of the polymer after its original assembly.

What is the effect of mixing a population of polymers assembled from a precursor pool of labeling probability  $p$  with a population of polymers assembled from an unlabeled precursor pool? Mixing of this sort (dilution of the polymer pool) is what happens in a biological system when a labeling experiment is performed: newly synthesized polymers from the labeled pool mix with polymers that were present before the experiment began. A key mathematical feature of MIDA is that the relationships among those polymeric species that contain labeled subunits (the internal pattern among isotopomers) are unchanged by dilution from an unlabeled population of polymers (Refs. 13–15; see CENTRAL FEATURES OF MIDA SUMMARIZED).

#### CENTRAL FEATURES OF MIDA SUMMARIZED

The first rule of MIDA is that there must be combinations possible in the molecule analyzed. At least two repeats of a probabilistically identical subunit must be present. Metabolic pathways involving other kinds of chemical transformations but no polymerization are therefore not amenable to the combinatorial approach. Polymers studied must also be analyzed intact, or with at least two subunits present, because the distribution of isotopomeric species carries the essential information. Any maneuver that reduces the population to monomeric homogeneity, such as combustion to carbon dioxide for isotope ratio measurements or hydrolysis to monomeric subunits before analysis, loses the combinatorial information and precludes application of MIDA.

The second rule of MIDA is that subpopulations of molecules must be distinguishable and quantifiable. Indeed, it is the variations within a population of assembled polymers that carry the information crucial for MIDA. The notion that there is a homogenous precursor pool and a uniform product pool is replaced by the notion of subpopulations of precursors (some A, some B) and subpopulations of products (of characteristic isotopomeric composition in quantifiable proportions). Any analytic modality must therefore be capable of discriminating among different polymeric subpopula-

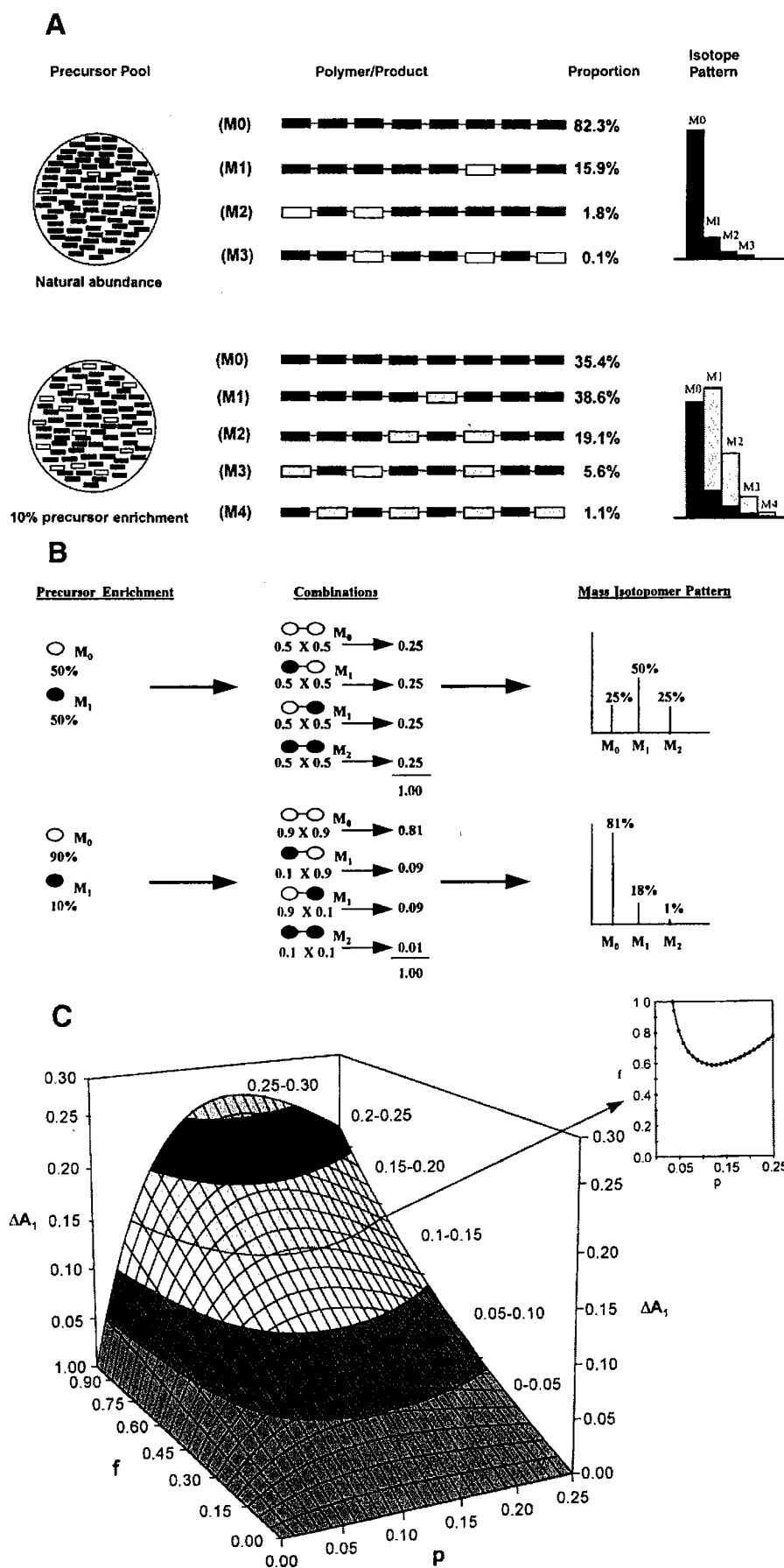


Fig. 1. Mass isotopomer distribution analysis (MIDA) principle. **A**: combinatorial probabilities determine the mass isotopomer pattern in polymers. In this simulation, natural abundance or 10% labeled pools of a subunit combine into a polymer of eight subunits. The population of each of these pools will contain a characteristic distribution of  $M_0$ ,  $M_1$ ,  $M_2$ , and so on, molecular species (mass isotopomers). These proportions can be represented as a frequency histogram of the mass isotopomer pattern in the polymer and can be measured by mass spectrometry. After correction for natural abundance, degree of enrichment of precursor pool can be calculated by comparing measured patterns of mass isotopomer abundances with those predicted from theoretical precursor pool enrichments. **B**: simple numerical example of MIDA principle. **C**: three-dimensional representation of change in fractional abundance ( $\Delta A_1$ ) of a particular mass isotopomer (in this case,  $M_1$  of methyl palmitate) as a function of  $p$  and  $f$ . A plane is extended at values of  $M_1 = 0.05, 0.10, 0.15, 0.20$  and  $0.25$ , demonstrating the family of solutions for all combinations of  $p$  and  $f$  that give this value of  $\Delta A_1$ . Inset: a 2-dimensional projection of  $p$  vs.  $f$  in the plane  $\Delta A_1 = 0.15$ . Note linearity of  $\Delta A_1$  vs.  $f$ .



tions (species) present within the population. This is why radioisotopic methods cannot be used: specific activity is measured from the total counts and total mass of material present, treated as a uniform population; and it is why average mass measurements by electrospray ionization-mass spectrometry also cannot be used: a "centroid" average mass collapses all of the population variability in the polymer pool into a single value.

The third essential concept underlying MIDA is that dilution of the monomeric (precursor) and polymeric (product) pools affects abundance distributions differently. Both sources of dilution can alter the relative proportion of polymeric species containing no labeled subunits vs. labeled subunits, but only dilution in the precursor pool can alter the internal quantitative relationships among labeled species. It is this differential effect on "amount" (proportion of the polymer population containing any labeled subunits) vs. "pattern" (relationships within the population of labeled polymers) that allows independent calculation of  $p$  and  $f$ , respectively.

The model just described involves some simplifying assumptions, which are discussed in THEORETICAL ISSUES FOR THE USE OF MIDA.

In addition to determining  $p$  directly, MIDA offers several operational advantages over previous isotopic techniques for measuring biosynthesis rates (Table 1). One analytic consequence of the need for intact combinations is that sophisticated mass spectrometric techniques must be used in cases in which biosynthesis of high-molecular-weight polymers, such as proteins or oligonucleotides, is being measured.

### Definitions

It is useful to define terms to avoid ambiguity. The following definitions will be used here.

**Isotopes.** Atoms with the same number of protons and hence of the same element but with different numbers of neutrons (e.g., H vs. D).

**Exact mass.** The mass calculated by summing the exact masses of all the isotopes in the formula of a molecule (e.g., 32.04847 for  $\text{CH}_3\text{NHD}$ ).

**Nominal mass.** The integer mass obtained by rounding the exact mass of a molecule.

**Isotopomers.** Isotopic isomers or species that have identical elemental compositions but are constitutionally and/or stereochemically isomeric because of isotopic substitution, as for  $\text{CH}_3\text{NH}_2$ ,  $\text{CH}_3\text{NHD}$ , and  $\text{CH}_2\text{DNH}_2$ .

**Isotopologues.** Isotopic homologues or molecular species that have identical elemental and chemical compositions but differ in isotopic content (e.g.,  $\text{CH}_3\text{NH}_2$  vs.  $\text{CH}_3\text{NHD}$  in the example above) (36). Isotopologues are defined by their isotopic composition; therefore, each isotopologue has a unique exact mass but may not have a unique structure. An isotopologue is usually comprised of a family of isotopic isomers (isotopomers) that differ by the location of the isotopes on the molecule (e.g.,  $\text{CH}_3\text{NHD}$  and  $\text{CH}_2\text{DNH}_2$  are the same isotopologue but are different isotopomers).

**Mass isotopomer.** A family of isotopic isomers that is grouped on the basis of nominal mass rather than

Table 1. *Some advantages and disadvantages of MIDA*

<i>Advantages</i>	
1.	No uncertainty regarding accuracy of true precursor pool estimate (e.g., concerns about compartmentalization of pools): the subunits that actually entered the polymer are the subunits that determine mass isotopomer abundance distributions.
2.	No need for isolation of putative precursor subunits.
3.	All parameters are derived from a single measurement (mass spectrometry of polymeric product).
4.	May be applied on multiple products simultaneously.
5.	Applicable in principle to all classes of polymers that can be stable isotopically labeled.
6.	Can be used in decay phase with high-mass isotopomers to achieve or confirm true pulse-chase conditions and determine turnover of polymer pool.
7.	Provides information about intracellular metabolite flow into and through precursor subunits (kinetics of the precursor).
8.	Can provide information about tissue source of polymers, subcellular pools, timing or location of biosynthetic events, metabolite channeling, or other characteristics of in vivo biosynthetic processes, on the basis of measurements of precursor pool enrichments.
9.	Does not typically require major perturbation of the physiological system studied, such as pool flooding, inhibition of synthesis or removal, exogenous labeling and reinjection, and the like.
10.	Not affected by isotope discrimination at the level of the precursor pool.
<i>Disadvantages</i>	
1.	May require relatively large precursor pool enrichments for some applications, which could alter metabolism due to non-tracer effects or be expensive to perform.
2.	Requires analytic accuracy, not just precision, because standard curves cannot routinely be generated (dilution in both precursor and product pools would need to be simulated).
3.	Requires analysis of intact polymers, or portions containing 2 or more repeats of a monomeric subunit, which can be analytically difficult, especially for large heteronuclear polymers (e.g., proteins, polynucleotides).

isotopic composition. A mass isotopomer may comprise molecules of different isotopic compositions, unlike an isotopologue (e.g.,  $\text{CH}_3\text{NHD}$ ,  $^{13}\text{CH}_3\text{NH}_2$ ,  $\text{CH}_3^{15}\text{NH}_2$  are part of the same mass isotopomer but are different isotopologues). In operational terms, a mass isotopomer is a family of isotopologues that are not resolved by a mass spectrometer. For quadrupole mass spectrometers, this typically means that mass isotopomers are families of isotopologues that share a nominal mass. Thus the isotopologues  $\text{CH}_3\text{NH}_2$  and  $\text{CH}_3\text{NHD}$  differ in nominal mass and are distinguished as being different mass isotopomers, but the isotopologues  $\text{CH}_3\text{NHD}$ ,  $\text{CH}_2\text{DNH}_2$ ,  $^{13}\text{CH}_3\text{NH}_2$ , and  $\text{CH}_3^{15}\text{NH}_2$  are all of the same nominal mass and hence are the same mass isotopomers. Each mass isotopomer is therefore typically composed of more than one isotopologue and has more than one exact mass. The distinction between isotopologues and mass isotopomers is useful in practice, because all individual isotopologues are not resolved using quadrupole mass spectrometers and may not be resolved even by using mass spectrometers that produce higher mass resolution, so that calculations from mass spectrometric data must be performed on the abundances of mass isotopomers rather than isotopologues. The mass isotopomer lowest in mass is repre-

sented as  $M_0$ ; for most organic molecules, this is the species containing all  $^{12}\text{C}$ ,  $^1\text{H}$ ,  $^{16}\text{O}$ ,  $^{14}\text{N}$ , and the like. Other mass isotopomers are distinguished by their mass differences from  $M_0$  ( $M_1$ ,  $M_2$ , etc.). For a given mass isotopomer, the location or position of isotopes within the molecule is not specified and may vary (i.e., "positional isotopomers" are not distinguished).

**Mass isotopomer pattern.** A histogram of the abundances of the mass isotopomers of a molecule. Traditionally, the pattern is presented as percent relative abundances, where all of the abundances are normalized to that of the most abundant mass isotopomer; the most abundant isotopomer is said to be 100%. The preferred form for applications involving probability analysis, such as MIDA, however, is proportion or fractional abundance, where the fraction that each species contributes to the total abundance is used (see ). The term isotope pattern is sometimes used in place of mass isotopomer pattern, although technically the former term applies only to the abundance pattern of isotopes in an element.

**Monoisotopic mass.** The exact mass of the molecular species that contains all  $^1\text{H}$ ,  $^{12}\text{C}$ ,  $^{14}\text{N}$ ,  $^{16}\text{O}$ ,  $^{32}\text{S}$ , and the like. For isotopologues composed of C, H, N, O, P, S, F, Cl, Br, and I, the isotopic composition of the isotopologue with the lowest mass is unique and unambiguous, because the most abundant isotopes of these elements are also the lowest in mass (23). The monoisotopic mass is abbreviated as  $m_0$ , and the masses of other mass isotopomers are identified by their mass differences from  $m_0$  ( $m_1$ ,  $m_2$ , etc.).

**Fractional abundances.** The abundances of individual isotopes (for elements) or mass isotopomers (for molecules) given as the fraction of the total abundance represented by that particular isotope or mass isotopomer. This is distinguished from relative abundance, wherein the most abundant species is given the value 100 and all other species are normalized relative to 100 and expressed as percent relative abundance. For a mass isotopomer  $M_x$

$$\text{fractional abundance of } M_x = A_x = \frac{\text{abundance } M_x}{\sum_{i=0}^n \text{abundance } M_i},$$

where 0 to  $n$  is the range of nominal masses relative to the lowest-mass ( $M_0$ ) mass isotopomer in which abundances occur.

$\Delta$ fractional abundance (enrichment or depletion) =

$$(A_x)_e - (A_x)_b = \left( \frac{\text{abundance } M_x}{\sum_{i=0}^n \text{abundance } M_i} \right)_e - \left( \frac{\text{abundance } M_x}{\sum_{i=0}^n \text{abundance } M_i} \right)_b$$

where subscript  $e$  refers to enriched and subscript  $b$  refers to baseline or natural abundance.

**Isotopically perturbed.** The state of an element or molecule that results from the explicit incorporation of an element or molecule with a distribution of isotopes that differs from the distribution found in nature

(Table 2), whether a naturally less abundant isotope is present in excess (enriched) or in deficit (depleted).

**Monomer.** A chemical unit that combines during the synthesis of a polymer and that is present two or more times in the polymer.

**Polymer.** A molecule synthesized from and containing two or more repeats of a monomer.

## CALCULATIONS

A tradition in mass spectrometric applications has often been to express quantitative results as relative abundances (each species normalized to the most abundant species, which is given the value 100). In contrast, fractional abundances or analogous expressions are generally preferable for MIDA, because the method is based on combinatorial probabilities, and probabilistic events are most directly represented as fractions of the total universe of choices possible.

## THEORETICAL ISSUES FOR USE OF MIDA

Every model is based on assumptions, which may or may not describe real biological systems accurately. It is useful to consider and ultimately to be able to evaluate or correct for potential deviations from the simple MIDA model. For all of the simulations performed here, calculations were carried out by use of the computer algorithms described in the APPENDIX.

### *Effect of Variations in Natural Abundance Values of the Isotopes of Elements*

The contribution to mass isotopomer distributions in a polymer from natural abundance isotopes of the elements has to be subtracted, or otherwise taken into account, for labeled subunits to be quantified (13, 15, 20, 22). One of the most obvious questions that is asked is whether the theoretical natural abundance values selected, especially for  $^{13}\text{C}$ , could have significant effects on the calculations. We have done calculations for several polymers while varying natural  $^{13}\text{C}$  fractional abundance between 1.08 and 1.11%, the range that might be present in biological carbon in mammals (6, 33), in the calculation algorithm. The results are shown for palmitate-methyl ester (Fig. 2). There is very little effect on calculated values of  $f$  when  $p > 0.03$ . The same is true for glucose, cholesterol (15), and other molecules (not shown). Thus variations in natural abundance values of  $^{13}\text{C}$  do not have an important effect on calculated parameters.

### *Effect of Isotope Discrimination*

Isotope discrimination or fractionation at the level of the precursor pool is not a problem within a biosynthetic model based on analysis of combinatorial probabilities (MIDA). The MIDA calculation reveals the isotope content of the subunits that actually entered a polymer, regardless of their relation to the isotope content of biochemical intermediates leading to these subunits. If, for example, there were a 10% discrimination against  $[^2\text{H}_3]$ leucine by leucyl-tRNA synthase or a 10% discrimination against  $[^2\text{H}_3]$ leucyl-tRNA by the ribosomal pro-

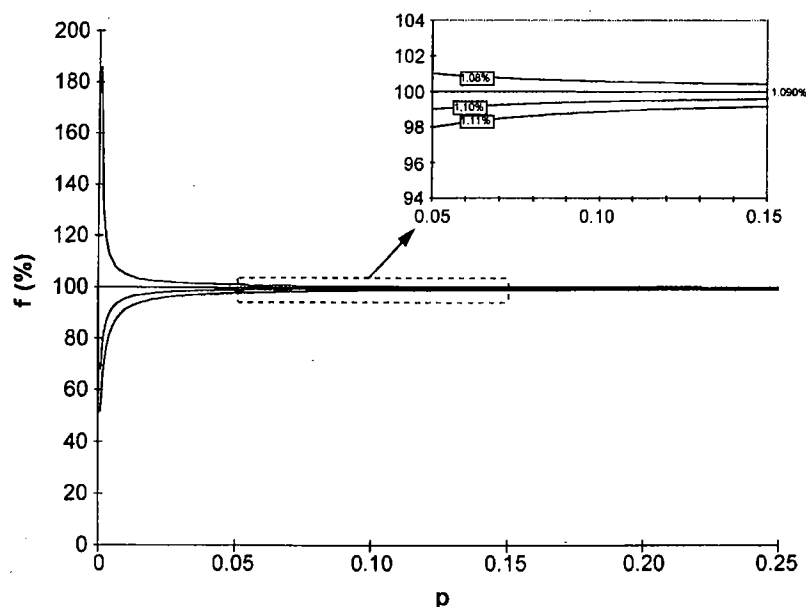


Fig. 2. Effect of varying the natural abundance  $^{13}\text{C}$  value on mass isotopomer abundances and MIDA calculations for methyl palmitate. Values of  $f$  are calculated using our standard algorithm ( $^{13}\text{C} = 1.09\%$ ), when sample in fact has natural abundance  $^{13}\text{C} = 1.08, 1.10$ , and  $1.11\%$ . Calculated values of  $f$  are shown as a function of  $p$  (calculated iteratively, at intervals of  $0.001$ ). *Inset*: blown-up region of graph between  $p = 0.05$  and  $0.15$ .

tein synthesis machinery, it would not affect the calculations of  $f$  or true  $p$ ; it is the  $^2\text{H}_3$  enrichment of the leucine subunits that actually entered the protein that determines the mass isotopomer abundance pattern. This pattern will reveal the true  $[^2\text{H}_3]$ leucine precursor subunit enrichment for biosynthesis, even if this value is different from the tRNA-leucine or free leucine enrichments, and the calculated fractional synthesis contribution would be correct. Measurements of  $p$ ,  $f$ , and the like would then be accurate with MIDA but not

if tRNA-leucine were used, in this example. Thus combinatorial probability analysis is unique among isotope kinetic approaches in that its validity is not altered by isotope discrimination during biosynthesis. In contrast, isotope effects on the metabolism of the polymer once synthesized (e.g., effects on clearance) will affect kinetic measurements, because the behavior of labeled polymers will not reflect that of the general pool.

#### Effect of Incorrect Value for Expected Number of Monomeric Subunits

Lee et al. (22) demonstrated elegantly that  $n$  (the number of precursor subunits actually present in a polymer) can be determined experimentally by using the same principles of probability analysis that are used for determining  $f$ . Instead of a reference table for  $p$  vs. mass isotopomer pattern at a known value of  $n$ , one can generate a reference table for  $n$  vs. mass isotopomer pattern at a known value of  $p$ . The true value of  $n$  can then be inferred from the experimental data. This technique is possible only when there exists an independent method for determining  $p$ ; the measurement of body water  $^2\text{H}$  enrichments during  $^2\text{H}_2\text{O}$  incorporation experiments represents a unique situation that permitted this application (22).

#### Calculation of $p$ and $f$ When the Complete Ion Spectrum is not Sampled

For any number of reasons, the mass spectrometrists may choose not to monitor all of the ions in a mass isotopomer envelope (e.g., for convenience, to maximize dwell time on the most abundant ions, or to avoid contaminating ions). An important property of combinatorial probabilities is that the calculation of  $p$  is not affected by the choice of ions selected for monitoring. The internal pattern and relationship among excess mass isotopomers are fixed and characteristic of  $p$  and  $n$  regardless of which particular masses are monitored.

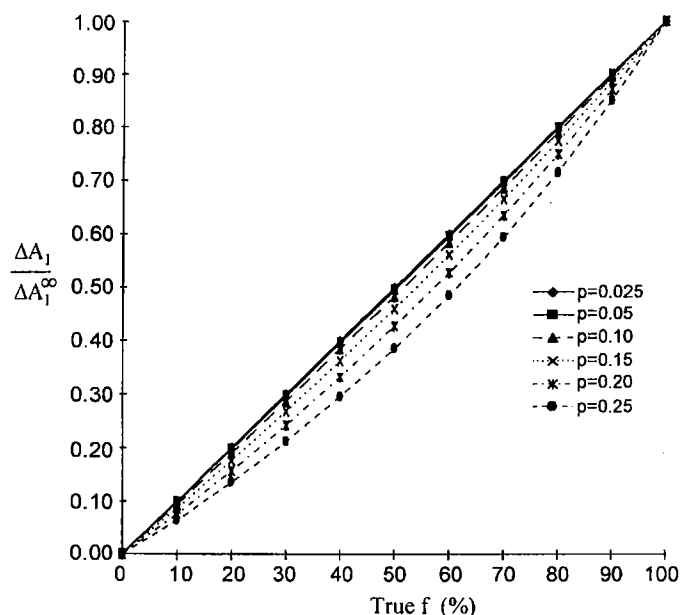


Fig. 3. Loss of linearity in the relationship between measured  $\Delta A_x$  and true  $f$ , especially at high values of  $p$ , if an incomplete ion spectrum is sampled. In this example, only ions  $M_0$  to  $M_2$  are measured in palmitate-methyl-ester at various values of  $p$ . The ratio  $\Delta A_x / \Delta A_1^\infty$  (measured excess  $M_1$ /asymptotic  $M_1$  values) should reflect  $f$  and rise linearly as  $f$  increases from 0 to 100% (e.g., see Fig. 1C). As  $p$  increases and the proportion of unmonitored ions rises, however, linearity is lost and true  $f$  is underestimated. Simulations were shown at data points of  $f$ .

As long as the appropriate equation is used for the masses under consideration, the choice of masses monitored will not influence calculation of  $p$ . Surprisingly, this is not the case for calculation of  $f$ , which is affected by incomplete ion spectrum sampling. This is because of a somewhat unexpected mathematical feature of mixtures of numerical distributions (e.g., populations of mass isotopomers): dilution is not linear when the proportion of the total population monitored is different in the natural abundance and enriched populations. As noted in the APPENDIX (and see Fig. 1C), the mathematical object of solving for  $p$  in step-wise calculation algorithms is to linearize the relationship between abundance of a particular mass isotopomer ( $A_x$ ) and the molar fraction of its associated molecular population in the mixture ( $f$ ), so that  $f$  can be solved algebraically from  $A_x$ . When different proportions of the total ion envelope are monitored for different populations, for example, as occurs when a high  $p$  generates high mass isotopomers that are not monitored, the linear relationship between any  $A_x$  and  $f$  (fractions of the mass isotopomer and molecule in the population, respectively) is lost (Fig. 3). Stated in intuitive terms, when higher masses are not monitored, a mole of isotopically enriched molecules will contribute fewer ions to the total spectrum sampled than a mole of natural abundance molecules. The molecular mixture is thereby weighted in favor of the more completely sampled molecular population (the unlabeled population), and a correction has to be made to put equal weight on each molecular population in the mixture.

A numerical example follows. If  $M_1$  theoretically represents 20% of the  $M_0$  through  $M_2$  ions in natural abundance molecules and 40% of the  $M_0$  through  $M_2$  ions at  $p = 0.10$ , but only 90% of the envelope is contained in  $M_0$  through  $M_2$  at  $p = 0.10$  compared with 100% in  $M_0$  through  $M_2$  for the natural abundance molecules, what would be the effect of mixing these populations and monitoring only  $M_0$  through  $M_2$ ? It should be apparent that the enriched population will not contribute the theoretical 40 ions of  $M_1$  for every 20 ions of  $M_1$  from unlabeled molecules in an equimolar mixture when only  $M_0$  through  $M_2$  are measured, but will contribute only 36 ( $= 40\% \times 90\%$  of the mole in the envelope monitored) for every 20 ( $= 20\% \times 100\%$  of the mole of natural abundance isotopomers). It would therefore be a mistake to use percentages of the ions monitored to represent percentages of the entire population when mixing populations with different percentages of ions monitored, because  $f$  will be systematically underestimated. The solutions to this confounding factor are straightforward: either monitor an essentially complete ion spectrum for the molecules under consideration or include a mathematical correction for unequal ion spectrum sampling in the calculation algorithm.

An algebraic correction is derived and presented in the APPENDIX (Eq. A9b) for instances of significantly incomplete ion spectrum monitoring. This equation corrects for the proportion of ions monitored at the measured value of  $p$  present relative to the proportion

monitored in unlabeled molecules. By use of this correction factor, mixtures of labeled and unlabeled molecules are again reduced to linear combinations of mass isotopomers, from which dilution of molecules can be calculated simply. This correction is generally extremely small and has no practical impact on most calculations (Fig. 4), because  $>98\%$  of the ions within an isotopomeric envelope are typically monitored for most labeled molecules. Failure to consider the effects of incomplete ion spectrum sampling can contribute to underestimation of values of  $f$  in special cases, however, such as very high values of  $p$ , if high masses are not monitored. The impact of incomplete ion spectrum monitoring has not to our knowledge been considered previously or corrected for in other MIDA calculation algorithms.

#### Effect of Analytic Artifacts—Fragment Ions

An important analytic concern is the effect on calculated parameters of fragment ions generated in the mass spectrometer source. The most common problem is the  $M_{-1}$  fragment, which can come from extraction of H from the parent ion. The effect of contaminating  $M_{-1}$  fragments can be simulated using the calculation algorithm described here (see APPENDIX). Contaminating  $M_{-1}$  fragments representing 0, 1, and 5% of ions were included in measurements of palmitate synthesis from acetyl-CoA (not shown). The contaminated fractional abundance distribution of mass isotopomers was calculated by overlaying the weighted intact and fragment ion distributions, and calculations of  $p$  and  $f$  were derived as though they had been generated as experimental data. Even at 5% contaminant contribution, calculated  $f$  was  $>99\%$ . Because the fragment contributes to natural abundance distributions as well as enriched distributions, its impact is substantially corrected through the subtraction of natural abundance values. It is nevertheless worthwhile to look for contaminating fragment ions in the mass spectrum.

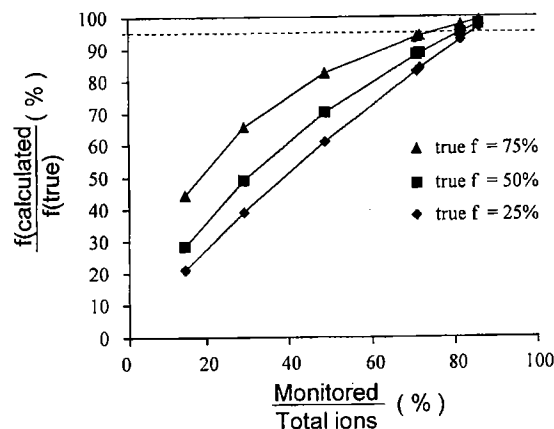


Fig. 4. Error in calculation of  $f$  resulting from incomplete ion spectrum monitoring of different degrees. Three different values for true  $f$  (25, 50, 75%) are shown. Note that calculated  $f$  remains  $\geq 95\%$  of actual value as long as  $\sim 85\%$  of the ion spectrum is monitored.

### *Effect of Isotopic Disequilibrium in the Precursor Pool*

*Evaluating the theoretical impact of isotopic disequilibrium.* A biosynthetic system may or may not exhibit isotopic equilibrium among different pools of a monomeric subunit from which a polymer is synthesized. Accordingly, there may not always be a single value for  $p$ . A physiologically relevant example of a biosynthetic system that does not necessarily have a single value for  $p$  is gluconeogenesis. There are actually two precursors comprising the gluconeogenic triose-phosphate pool, dihydroxyacetone phosphate (DHAP) and glyceraldehyde-3-phosphate. Isotopic equilibrium between the triose-phosphates may not always be complete. One can simulate the effect of various degrees of isotopic disequilibrium between DHAP and glyceraldehyde-3-phosphate and determine the extent to which calculated values of  $p$  and  $f$  would be distorted if the standard MIDA reference table were applied to the  $\Delta$ fractional abundances generated (27). If we allow the average  $p$  to range between 0.05 and 0.15 and vary  $p$  in *pool 2* from 1.0 to 2.0 times the *pool 1* value, the consequences for the  $f$  can be calculated. When  $p$  in *pool 1* =  $p$  in *pool 2* (i.e., when isotopic equilibrium is present), the calculated value of  $f$  is of course exactly as expected from the MIDA tables (100%), whatever the value of  $p$ . When  $p$  in *pool 2* is double that of *pool 1*, calculated  $f$  is ~12% higher than the true value for all values of  $p$ ; when *pool 2* is made 50% above *pool 1*,  $f$  is within 4% of the true value. An interesting general mathematical result to emerge from this analysis is that  $f$  is always overestimated (>100% the actual value) if isotopic disequilibrium exists within precursor subunit pools (27). The practical implication for measuring gluconeogenesis in particular is that isotopic disequilibrium in the triose-phosphate precursor pool is in principle unlikely to represent a major problem, because the MIDA calculations will work well unless DHAP and glyceraldehyde-3-phosphate are differentially enriched by a factor of >2 (27). Similar calculations can be applied to other polymers of interest.

*Correcting for documented isotopic disequilibrium.* If one is not satisfied with theoretical arguments discounting the importance of isotopic disequilibrium within a precursor pool, it is possible to modify the calculation algorithm to incorporate deviations from isotopic equilibrium within precursor pools. Theoretical tables can be generated by modifying the algorithm described above, if the degree of isotopic disequilibrium is measurable. In the case of gluconeogenesis, for example, mass spectrometric fragmentation of the molecule into "top" (C-1 to C-3) and "bottom" (C-4 to C-6) halves can reveal whether labeling was equal in DHAP and glyceraldehyde-3-phosphate, respectively (26). If enrichments differ by a certain proportion, then this value can be incorporated into the probability calculations to generate an appropriate, individualized standard curve that adjusts asymptotic values appropriately. Thus individualized standard curves can be generated for each experiment on the basis of the observed degree of isotopic disequilibrium within precursor pools, if the

latter is significant and can be determined experimentally. Because the mathematical approach that we have described (see APPENDIX and Refs. 13–15) is based on empirical relationships between derived values ( $\Delta$ fractional abundances) rather than expressions of the pure binomial or multinomial expansion, deviations from the simple combinatorial probability model can be accounted for relatively easily and without compromising mathematical rigor.

### *Effect of Isotopic Inhomogeneity in the Precursor Pool (e.g., More Than One Biosynthetic Site, an Isotopic Gradient Across a Tissue, or Time Variations in $p$ )*

A potentially more important deviation from the simple MIDA model is if the polymer is made in more than one anatomic location and  $p$  is not equal in each site, or if the value of  $p$  changes over time. This situation differs from the situation of isotopic equilibrium within a precursor subunit pool (see *Effect of Isotopic Disequilibrium in the Precursor Pool*). Each of the polymer populations synthesized has a single precursor enrichment, but there is more than one pool of polymers present, whereas in the former case, each polymer molecule has more than one precursor enrichment but there is only a single pool of polymers. Examples of more than one anatomic location for biosynthesis might include extrahepatic and hepatic gluconeogenesis or cholesterologenesis, or labeling gradients within a precursor pool across a tissue. The same considerations would apply if  $p$  changed over time during an experiment; polymers made at different times would have different mass isotopomer patterns.

The consequence of these scenarios is that, instead of a single binomial distribution, there will be a mixture of distributions from the different values of  $p$  in newly synthesized polymers. This mixture of distributions itself mixes with, and needs to be distinguished from, the natural abundance distribution that represents old molecules. Any attempt to model biosynthesis as two populations or two distributions, enriched and unenriched, will not be rigorously correct mathematically but will instead be an approximation, because binomial or multinomial expansions cannot themselves be averaged (i.e., are not linear) (21). The  $M_2$  isotopomer changes approximately as the power of 2 for changes in  $p$ , the  $M_3$  as the power of 3, etc., so that different expansions cannot be combined and averaged as if they were linear.

The practical questions are, how much does this matter? what impact on estimated parameters will there be if  $p$  is inconstant in time or space? and can it be identified or corrected for when present? A simulation for gluconeogenesis in two tissues or at two time points with different  $p$  values has been presented elsewhere (26):  $f$  remains  $\geq 0.8$ –0.85, even when *pool 2* enrichment is 2–3 times *pool 1*. Another example is a gradient in precursor pool enrichment across a tissue (Fig. 5A). If one models 10 pools contributing equally to gluconeogenesis with a labeling gradient that spans an approximately twofold range [e.g., 0.105–0.195 molar excess (ME)], the consequence is minor (underestimation of

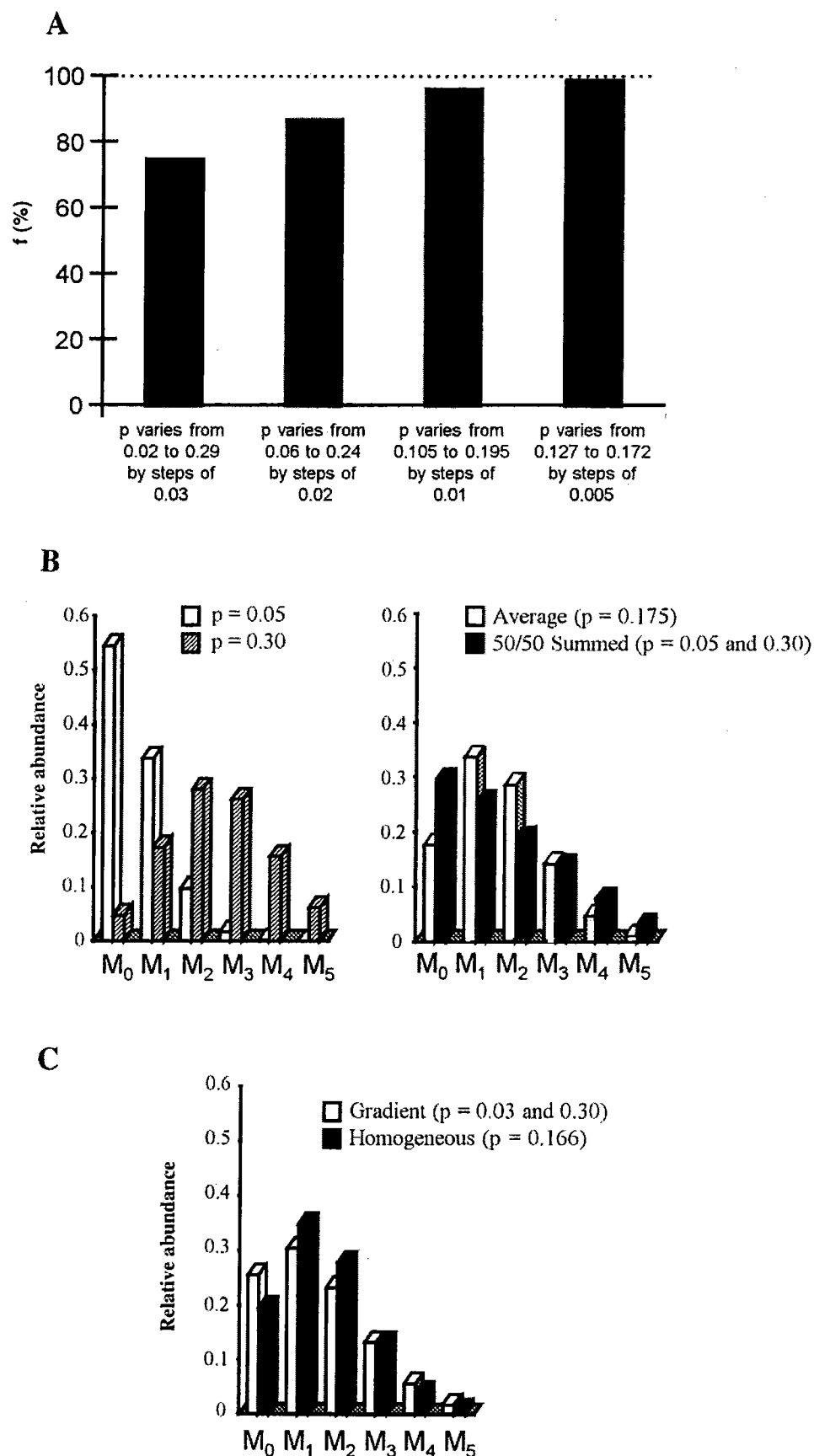


Fig. 5. Effects of isotopic inhomogeneity in the precursor pool on calculated values of  $p$  and  $f$ . *A*: effect on gluconeogenesis of an isotopic gradient within a tissue. Ten sites are simulated to contribute equally to gluconeogenesis, with different degrees of isotopic gradient in the region around  $p = 0.15$  (from 0.125–0.175 up to 0.01–0.29, at the intervals stated). *B*: use of high-mass isotopomer analysis to identify inhomogeneity in the precursor pool, or in this case, two precursor pools. Isotopomeric envelopes resulting in palmitate-methyl-ester if two biosynthetic sites contribute equally, one at  $p = 0.05$  and one at  $p = 0.30$ , are shown (*left*). The envelope from the summed distribution is then compared with the envelope expected for  $p = 0.175$  (the average value) (*right*). *C*: use of high-mass isotopomer analysis to identify inhomogeneity in the precursor pool, in this case, an isotopic gradient in the precursor pool. A labeling gradient from 0.03 to 0.30 for  $p$  is compared with a homogenous pool at the value of  $p$  (0.166) calculated from the observed  $M_2$ -to- $M_1$  ratio.

true values of  $f$  by a factor of  $<5\%$ , Fig. 5A). Even for a fourfold gradient (0.06–0.24 ME),  $f$  is only underestimated by  $\sim 15\%$ ; i.e., if the actual value of  $f = 0.50$ , measured  $f$  will be 0.425. Only at very large gradients (e.g., 15-fold, from 0.02 to 0.29 ME) is even a 25% underestimation observed. An analogous situation can be simulated for lipogenesis, with an isotopic gradient of acetyl-CoA across a lipogenic tissue such as liver. If 10 pools contribute equally to lipogenesis, with a gradient from 0.03 to 0.30 ME (at intervals of 0.03),  $f$  is 87.3% instead of 100% (i.e., underestimated by 12.7%). If a gradient from 0.02 to 0.10 ME is simulated, with 100 pools contributing equally to lipogenesis, the value of  $f$  calculated by MIDA by use of  $M_1$  and  $M_2$  isotopomers is 90.1%. Thus inconstancy in  $p$  over time or space means that the simple binomial model becomes an approximation rather than an exact description of biosynthesis, but the practical impact varies according to physiological conditions and typically is fairly small. An investigator is able to evaluate the likelihood of significant error and the practical acceptability of this degree of error by performing a simulation of this type.

It is also possible to identify, and even correct for, isotopic gradients or variations in precursor pool enrichments by application of combinatorial principles. If a large isotopic gradient is present or two precursors of different enrichment contributed to biosynthesis, there will be a divergence between the isotopomer pattern in high-mass (multiply labeled) vs. low-mass (single- and double-labeled) polymers (Fig. 5, B and C). If a gradient exists within the precursor pool, high-mass species will be produced that would otherwise never be observed if the average value of  $p$  were uniformly present. The pattern of higher-mass isotopomers predicted by analyzing lower masses will not be observed; similarly, the pattern of low-mass isotopomers predicted by the higher masses will not be met (Fig. 5, B and C).

The occurrence of "inappropriate" multiply labeled species relative to the pattern among the less-labeled species (Fig. 5, B and C) can therefore be used diagnostically to confirm or exclude the existence of an isotopic gradient. One simple approach is to monitor higher-mass isotopomers and compare calculated values of  $p$  from higher- vs. lower-mass relationships. In the case of a labeling gradient across a lipogenic tissue spanning 0.03 to 0.30 (Fig. 5C), for example, the pattern of excess  $M_3/M_2$  isotopomers in palmitate indicates  $p = 0.213$ ; the ratio in  $M_4/M_3$  indicates  $p = 0.226$ , whereas  $M_2/M_1$  isotopomers reveal a lower  $p$  (0.166). The divergence is even greater for simple two-precursor-pool systems (Fig. 5B). If there were an equal mixture of a palmitate population derived from 0.30 and a 0.50 value of  $p$ , analysis of  $M_1$  and  $M_2$  would reveal  $p = 0.173$ , whereas analysis of  $M_3$  and  $M_4$  would indicate  $p = 0.29$ . In contrast, if a single, homogenous precursor pool is present, calculated values of  $p$  are identical whatever masses are used for its calculation (15). The finding of different  $p$  values by analysis of high- vs. low-mass isotopomer patterns represents a technique for identifying the presence of an isotopic gradient.

### *Conversion of Fractional Synthesis Values into Chemical Fluxes: Combining MIDA Calculations with Administration of Exogenous Stable Isotope-Labeled Polymers*

Expression of synthesis as rates in chemical units (mass/time) requires an estimate of the turnover of the polymer pool being sampled in addition to the fraction of the polymer pool that came from endogenous synthesis during the time period studied. Probability considerations demonstrate that high-mass isotopomers are uniquely useful in the labeled-polymer decay phase, because problems from persistent isotope incorporation are avoided for multiply labeled species even if pulse/chase conditions do not exist (i.e., if precursor subunits continue to contain a low level of labeling because of isotope recycling or slow turnover of the precursor pool). This application of MIDA has been discussed in detail previously (15, 25). The turnover of the polymer can also be measured by analyzing the rate of rise toward plateau during the label incorporation phase, as discussed elsewhere (16). Alternatively, exogenously labeled polymers can be administered (to measure turnover by dilution) concurrently with a biosynthetic incorporation experiment. For this last approach, however, potential interference by the exogenous label with the isotope incorporated via biosynthesis has to be accounted for.

For example, it is useful to measure the plasma glucose turnover concurrently with fractional gluconeogenesis to determine the absolute rate of gluconeogenesis. If  $[1\text{-}^2\text{H}]\text{glucose}$  or  $[6,6\text{-}^2\text{H}_2]\text{glucose}$  is used for turnover, labeled species have the same nominal mass as the isotopomers analyzed for MIDA calculations ( $M_1$  and  $M_2$ ). Because quadrupole mass spectrometers do not have sufficient resolving power to distinguish between  $^2\text{H}$  and  $^{13}\text{C}$  exact masses, a technique for determining the contributions from  $[^2\text{H}]\text{glucose}$  vs.  $[^{13}\text{C}]\text{triose-phosphate}$  is required. This can be achieved by analyzing a derivative that is stripped of the labeled hydrogen (e.g., aldonitrile-pentaacetate or saccharic acid derivatives from which position 1 and 1,6 hydrogens are removed, respectively) in addition to a derivative that contains both inputs (e.g., pentaacetate). The MIDA calculation on the derivative stripped of  $^2\text{H}$  is routine. The problem, however, is how to "subtract" or correct the  $^{13}\text{C}$  contribution from the combined spectrum to establish the  $^2\text{H}$  labeling, by difference.

A calculation algorithm can be used to correct for the underlying isotopomeric distribution from incorporation of  $^{13}\text{C}$  gluconeogenic precursors, for example, to measure  $[^2\text{H}]\text{glucose}$  enrichment (7). The glucose molecules present during a simultaneous measurement of fractional gluconeogenesis and glucose turnover consist of a mixture of three populations: gluconeogenic product molecules arising from the labeled triose-phosphate precursor pool, labeled glucose molecules infused exogenously as tracer, and unlabeled molecules with a natural abundance distribution. The key is that the isotopomeric distribution of each of these components is known. The infused tracer ( $[6,6\text{-}^2\text{H}_2]\text{-}$  or  $[1\text{-}^2\text{H}]\text{glu-}$

cose) has an isotopomer distribution that is easy to calculate; natural abundance glucose also has a known distribution; and the gluconeogenic population has a distribution of isotopomers that is a function of  $p$  and that is measurable from the deuterium-stripped derivative. The distributions from each of these three components of the mixture can therefore simply be added to construct a theoretical standard curve, simulating the effect of adding  $^2\text{H}$ -labeled glucose to mixtures of the other two populations at the measured  $f$ . If gluconeogenesis  $f = 0.33$  and  $p = 0.15$ , for example, the calculation consists of adding 0.33 times the abundance of each isotopomer from gluconeogenesis at  $p = 0.15$ , then  $(0.67 - z)$  times the abundance of each isotopomer from natural abundance glucose, and  $z$  times the abundance of each isotopomer from the  $[^2\text{H}]$ glucose, where  $z$  is the fraction of  $^2\text{H}$ -labeled glucose added to generate the theoretical standard curve. Two or more values of  $z$  are simulated to construct a linear standard curve, wherein the isotopomers of interest are plotted against  $z$  to generate a slope and intercept for calculation of  $^2\text{H}$  enrichment and dilution in the intact molecule (Fig. 6). This algorithm must be applied separately for each time point sampled, because each sample will have a unique  $p$  and  $f$  (8, 36), and thus a unique slope and intercept for the standard curve.

#### *Assessing the Sensitivity of MIDA-Calculated Biosynthetic Parameters to Measurement Error*

To plan isotopic experiments and evaluate results, it is useful to know how sensitive the derived parameters are to analytic imprecision or inaccuracy. Error-sensitivity analysis can be performed by using the MIDA calculation algorithm presented here. An example with gluconeogenesis has been presented previously (26). This analysis reveals that an analytic coefficient of variation of 1% in estimates of the ratio of  $\Delta A_2$  to  $\Delta A_1$  at  $p = 0.15$  alters estimates of  $p$  by  $\sim 2\%$  and  $f$  by 1.6%. In contrast, at  $p = \sim 0.05$ , an analytic coefficient of variation of 1% alters estimates of  $p$  by 5.2% and  $f$  by 4.9% (26). It is apparent from this analysis that the experimenter is better advised to aim for  $p$  in the range of 0.15 rather than 0.05 to reduce sensitivity of final results to analytic error. Because analytic precision of

better than  $\pm 0.02$  mole percent excess (MPE) can be attained by using multiple replicate analyses, several time points,  $\alpha$ - and  $\beta$ -anomers of glucose, and other analytic strategies (26), this analysis also reaffirms the reproducibility of estimates by MIDA for applications such as gluconeogenesis, when  $p$  is in the range of 0.15. The interaction between enrichment achieved in a labeling experiment (i.e., signal) and sensitivity to analytic error is discussed further in ANALYTIC AND EXPERIMENTAL DESIGN CALCULATIONS.

#### ALTERNATIVE CALCULATION ALGORITHMS

MIDA as we now understand it was developed from 1990 to 1992 by Hellerstein and co-workers (13–15, 17) and Kelleher et al. (20), working independently. MIDA is defined by its purpose, its method, and its calculations. The purpose or field of MIDA is to measure the synthesis of biological polymers, the isotopic labeling of the monomeric precursor pool from which the polymers were assembled, and related kinetic parameters. The method involves measurement and analysis of mass isotopomer abundance distributions in intact polymers according to a combinatorial probability model, after introduction of a stable isotopically labeled monomeric subunit. The calculation approach uses the binomial or multinomial expansion as a basis for interpreting incorporation of isotopically labeled repeating subunits in the intact polymer and for inferring dilution in the monomeric and polymeric (precursor and product) pools.

It is in the area of a calculation algorithm that modifications have been presented (1, 5, 22, 32) since the original approaches described by Hellerstein and co-workers (13–15, 17) and Kelleher et al. (21). All the calculation algorithms presented so far, however, share fundamental assumptions and postulate a common model of biosynthesis. All postulate a combinatorial (binomial/multinomial) precursor-product biosynthetic model, and all postulate two confounding factors that may then modify the simple binomial/multinomial distributions: first, the natural abundance isotopes in the molecule, which interact with label-derived isotopomers; and second, the fact that two sources of dilution exist in biosynthetic systems (in the product pool as well as the precursor pool) and that these two sources of dilution influence isotopomer abundances differently in the product.

Only two general solutions have been proposed for each of the problems just noted. The influence of natural abundance isotopes on label distributions is a strictly computational problem. The solution has been either to create a computational model that incorporates all sources of isotope from both labeled precursor and natural abundance isotopes, and thus does not conform to a simple binomial distribution (Hellerstein and Neese, Ref. 15, and Kelleher et al., Ref. 20), or to transform the data in a way that removes the influence of natural abundance isotopes and restores a pure binomial distribution from the labeled precursor (Lee et al., Ref. 22, and Chinkes et al., Ref. 5). The problem of two sources of dilution, in contrast, reflects a biological issue. Again, there have been two computational solutions proposed. Most methods (Hellerstein, Lee, and

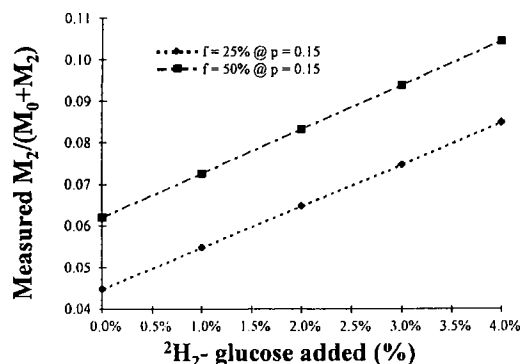


Fig. 6. Standard curve for calculation of  $[6,6-^2\text{H}_2]$ glucose enrichment in the presence of an underlying  $^{13}\text{C}$  combinatorial distribution from gluconeogenesis. In this example, the value for  $p$  is 0.15.



Chinkes) have used a step-wise approach. Ratios among natural abundance-corrected terms are first computed. Use of internal ratios of isotopomers in the polymer sample analyzed removes the effect of varying product dilution, because all isotopomers sampled in the labeled molecules are equally diluted by natural abundance molecules, so that precursor pool dilution can be calculated independently. Once this unknown ( $p$ ) is solved, the second unknown ( $f$ ) can be solved algebraically (Fig. 1C). Alternatively, the two sources of dilution (Fig. 1C) can be solved simultaneously, by best fit (nonlinear regression analysis) of multiple solution sets for the two unknowns taken together (21).

The most important point is that all the calculation algorithms presented to date give essentially identical results. Thus data of Byerley et al. (4), from  $^2\text{H}_2\text{O}$  incorporation into cellular cholesterol in cultured cells, give identical values for  $p$  and  $f$  whether calculated by the method of Lee (22) or by our approach (unpublished results). Identical results are obtained when data are analyzed by the approach of Kelleher et al. (21) and ours (unpublished observations; and T. Masterson, personal communication, April 1994). Chinkes et al. (5) also compared calculation algorithms using simulated data and concluded that results were essentially identical.

It should be noted that all these approaches are of roughly equal computational complexity. They all require a computer program<sup>1</sup>, to calculate abundances and to generate either a model or algebraic correction factors, and a specific software package that has to be applied separately for each molecule analyzed.

#### ANALYTIC AND EXPERIMENTAL DESIGN CONSIDERATIONS

The single most difficult problem facing the use of MIDA at present, both in theory and in practice, relates to quantitative accuracy of measurements, i.e., the analytic performance of mass spectrometers. This problem is widely recognized by workers in the field but has only rarely been noted in the literature (3, 10, 11, 23, 24, 26, 30). MIDA is based on analysis of numerical distributions in the context of a model of combinatorial probabilities. If the instrument generates inaccurate numbers, measured distributions will no longer reflect the actual isotopomeric distributions present. The actual effect on kinetic estimates ( $p$ ,  $f$ ) will depend on the nature and extent of the experimental inaccuracy. How then does one apply equations that are based on a model of combinatorial probabilities if the numbers do not fit the actual distributions generated? The best solution would be if the experimenter understood the cause and exact nature of instrumental inaccuracy: if one knew that the explanation was, for example, inadequate resolution of adjacent masses by the mass analyzer, a suitable correction algorithm might be applied. This is in essence what is done with liquid scintillation counting of radioisotopes, wherein  $^3\text{H}$  and

$^{14}\text{C}$  spillover is accounted for, so that each isotope can be independently measured. Unfortunately, the analytic basis of quantitative inaccuracy by current mass spectrometers is not understood in a sufficiently definitive way at present to allow simple correction (see *Strategies for Evaluating Quantitative Instrument Performance and Data Acceptability*). Another reasonable approach is to use standard curves, as one does when measuring dilution of an exogenous labeled product. There are problems also that make this difficult for biosynthetic MIDA methods, however. These we will discuss.

Several questions concerning mass spectrometric quantitative inaccuracy need to be addressed. We will not review these issues extensively here but will note some practical implications of each.

#### *The Nature of Quantitative Inaccuracy in Mass Spectrometric Measurement of Isotope Ratios*

Surprisingly little literature exists concerning the mass spectrometric causes of deviations between expected and measured abundances of mass isotopomers in organic analytes (10, 11, 30). The abundance of one mass relative to another might be overestimated by a fixed proportion, by a fixed amount, or by random error. Each type of error would have different implications for attempted correction algorithms. One also needs to know whether instruments drift over time, so that standards analyzed near to a sample in time will reveal the existence of error in the sample, or whether error occurs in an erratic, unpredictable way. Unfortunately, few data and no consensus exist on these questions, although several mechanistic possibilities can be considered.

One potential cause of inaccuracy is incomplete resolution of adjacent ions in the ion envelope (peak tailing), resulting in contamination of adjacent mass channels. If the mass analyzer is operating at the limit of its mass-resolving capacity, the degree of misidentification of ions due to ion scattering and peak tailing (24) could vary from run to run. According to this explanation, mass analyzers that achieve significantly better resolution (e.g., magnetic sector compared with quadrupole mass analyzers) would be predicted to exhibit better accuracy. This prediction has not yet been systematically tested. Another prediction is that higher abundances in adjacent mass isotopomers should worsen resolution (increase peak tailing), resulting in worse concentration sensitivity of fractional isotopomer abundances. Abundance sensitivity (24), the observation that the most abundant ion in an envelope tends to be underestimated quantitatively, has been identified as a problem in the field of isotopic analysis of elements by use of isotope ratio mass spectrometers. The physical explanation and methods for instrumental correction continue to be debated, however. A second possible cause of inaccuracy is nonlinearity of the detector response at different abundances or for different ions. If the detector output at each mass does not faithfully and consistently reflect the number of ions that reach it, numerical distributions will be skewed (10). Detector

<sup>1</sup> A computer program for calculating theoretical tables and spreadsheets is available and will be sent by the authors on request.

Table 2. *Instrument factors that might influence measured abundances of mass isotopomers*

Cause	Effect	Direction of			Dependent on Amount of Sample?	Instrument Dependence		
		Effect for $M_0$	Effect for $M_1$	Effect for $M_2$		TOF	Quadrupole	Sector
Bimolecular ion chemistry	Addition of H	—	+++	+	Yes	Yes	Yes	Yes
Bimolecular ion chemistry	Addition of 2 H units	—	NA	++	Yes	Yes	Yes	Yes
Isotope effects during fragmentation	Misrepresentation of relative abundances	Var	Var	Var	No	Yes	Yes	Yes
Incomplete resolution (peak tailing)	Misidentification of adjacent ions	---	+++	+	No	Yes	Yes	No
Decreasing ion transmission efficiency with mass of ion	Mass-dependent MS response	+	—	--	No	No	Yes	Some
Nonlinear signal amplification	Nonlinear signal output	--	++	+++	Yes	Yes	Yes	Yes
Saturation of multiplier	$M_0$ underrepresented	--	++	++	Yes	Yes	Yes	Yes
Velocity dependence of multiplier output	Mass-dependent MS response	+	—	--	No	Yes	Yes	Yes
Integration: baseline too high	Less abundant ions underrepresented	+	—	---	No	UN	UN	UN
Integration: baseline too low	Less abundant ions overrepresented	—	+	+++	No	UN	UN	UN
"Abundance sensitivity"	Underrepresentation of more abundant ions	--	—	++	Yes	UN	Yes	UN

TOF, time-of-flight mass analyzer; sector, magnetic sector instrument; NA, not applicable; +, increases (slightly); ++, increases (moderately); +++, increases (strongly); —, decreases (slightly); --, decreases (moderately); ---, decreases (strongly); Var, variable; UN, unknown.

nonlinearity might be correctable by use of suitable standards, if these could be synthesized, or by use of improved multiplier-detectors. Detector nonlinearity would also predispose to abundance-sensitivity effects: if abundances of different isotopomers span a large dynamic range, then the slope of detector output vs. injected material will be different for each. A third possibility is the occurrence of chemistry in the ion source (10, 11, 30). Chemical reactions such as hydrogen abstraction or addition could alter mass isotopomer abundances. This problem could be addressed by using derivatives that are unlikely to undergo these reactions (such as fluorinated molecules lacking exchangeable hydrogens) or by use of different ionization conditions (e.g., metastable atom bombardment). This problem might also explain concentration effects on isotope ratios (increased concentrations in the ion source lead to more chemical interactions, as postulated in Refs. 10, 11, and 30).

Regardless of the physical mechanisms involved, some empirical observations may be helpful operationally. First, concentration effects (relative abundances of isotopomers varying as a function of the total amount of material injected onto the mass spectrometer) are a problem for most molecules and tend to be more pronounced the greater the dynamic range among the masses monitored. With glucose-pentaacetate, for example, in natural abundance samples the ratio of the  $M_0$  (0.8396) isotopomer to the  $M_1$  (0.1348) and  $M_2$  (0.0256) isotopomer abundances spans almost two orders of magnitude. In contrast, at higher values of precursor enrichment (e.g.,  $p = 0.24$ ), the ion envelope is more evenly spread out ( $M_0 = 0.3060$ ,  $M_1 = 0.4400$  and  $M_2 = 0.2550$ , for a span of less than twofold), and observed concentration sensitivity is in fact less of a problem (Neese, R. A., R. Bandsma, and M. K. Hellerstein, unpublished observations). Second, the highest abundance isotopomers tend to be relatively underesti-

mated as total ion abundance increases (as has also been observed with inorganic elemental analyses using isotope ratio mass spectrometers, Ref. 24). Third, and counterintuitively, masses with very low baseline isotope abundances (i.e., higher masses with few or no natural isotope abundances) can be analytically undesirable when they have to be compared quantitatively with higher abundance masses. Although it may seem attractive to avoid having to subtract baseline values, the extremely large dynamic range maximizes the relative concentration sensitivity of different isotopomers and may thereby lead to concentration-sensitivity for isotopomer abundances.

In summary, a number of analytic factors influence the relative abundances of mass isotopomers measured in an envelope (Table 2). These factors range from chemical events in the ion source (abstraction of H) to performance of the mass analyzer (ion transmission efficiency, mass resolution, peak tailing), characteristics of the ion detector (velocity dependence of multiplier, nonlinearity due to threshold or saturation effects), or accuracy of the integration software (baseline value that is subtracted, peak-fitting algorithm used). Until basic mass spectrometry research identifies the causes of quantitative inaccuracy, it will not be possible to correct post hoc in a definitive manner for substantial deviations from expected isotopomer abundances. Other strategies are therefore required. The most important of these in practice involve assessment of instrument performance and analytic techniques for prevention of inaccuracy.

#### *Strategies for Evaluating Quantitative Instrument Performance and Data Acceptability*

In the absence of methods for salvaging inaccurate analyses, the practitioner can establish criteria for acceptable accuracy and reject data that fail to meet

these criteria. Two general approaches can be used for evaluating instrument accuracy and data acceptability.

*Measurements on natural abundance standards.* The theoretically expected mass isotopomer abundances can be calculated for any natural abundance molecule or ion fragment of known chemical composition (see APPENDIX). The impact of variations in natural isotope abundances is extremely small (Fig. 2), so that natural abundance distributions should reliably predict measurements from standards. The simplest test of instrument accuracy is therefore how closely the measured isotopomer abundances in a natural abundance sample match the theoretically expected values. Minimum accuracy criteria for data acceptability can be empirically established on the basis of statistical considerations (e.g., all isotopomers must be within 2% of their theoretical natural abundance values for the data to be included; thus 0.0025 for an  $M_1$  of palmitate-methyl ester or 0.0005 from  $M_2$  of palmitate-methyl-ester or glucose pentaacetate), taking into account the degree of accuracy required for a particular application, which in turn is influenced by the isotopomer enrichments achieved in the experiment. Obviously, accuracy needs to be at its best if the actual isotope enrichments present are low. One can formalize these criteria by simulating the effect of different types of errors, e.g., constant fraction or constant amount errors on calculated parameters (see *Achievement of adequate enrichments of masses of interest*).

Comparison of baseline measurements to theoretical abundances represents a simple strategy for establishing whether mass resolution, nonlinearity of detector response, or ion chemistry is a significant problem in a particular analysis. The main caveat is that analysis of natural abundance molecules cannot ensure that linearity is maintained as the relative abundances of different ions change, i.e., that accuracy will be maintained for labeled, biological samples. The latter may require use of labeled standards. Also, it should be noted that performance for mass isotopomers with essentially zero natural abundance values (e.g.,  $M_4$  glucose) cannot be evaluated by this method, which makes it more difficult to assess quality of data if these masses are subsequently monitored in labeled molecules (32).

*Use of isotopically enriched mass standards to establish linearity of response for different isotopomers.* One can use labeled standards to assess linearity and quantitative accuracy at various masses.  $[1-^2\text{H}_1]$ - and  $[6,6-^2\text{H}_2]$ glucose or  $[1-^{13}\text{C}]$ - and  $[1,2-^{13}\text{C}_2]$ palmitate, for example, can be purchased and mixed with natural abundance molecules to generate mass standards. After correction for the isotopomeric envelopes actually present in labeled molecules (see APPENDIX), standard curves representing different values of  $p$  and  $f$  can be analyzed. Problems with this approach include the fact that standards are themselves not 100% isotopically enriched, so one has to use an instrumental measurement of enrichment at some point to evaluate instrument performance, which can lead to circularities of logic, and the fact that accuracy of pipetting, completeness and comparability of derivitization, and reproduc-

ibility of the injector can influence the standard curves observed. Moreover, this approach does not provide a definitive method of correcting for inaccuracies of abundance measurements, but only of evaluating whether inaccuracy is present and severe enough to invalidate results. Nevertheless, this approach is useful for evaluating quantitative linearity over a range of isotopomer abundances (3, 28) and is particularly useful when a mass isotopomer being monitored has insufficient natural abundances to be measurable in baseline samples.

#### *Guidelines for Optimizing Quantitative Mass Spectrometric Analyses*

There are some useful experimental guidelines that have proven helpful in preventing instrument inaccuracy and reducing the impact of instrument performance on kinetic parameters with the use of MIDA (26, 30).

*Attention to concentration effects.* Publications that use MIDA should show the concordance of baseline isotopomer ratios with expected abundances (or show some other index of accuracy) and should demonstrate that concentration effects were considered and avoided. Failure to report these data makes technical evaluation of results difficult or impossible.

*Achievement of adequate enrichments of masses of interest.* A second factor that is to some extent under the control of the investigator is the enrichments achieved in mass isotopomers of interest. Improving enrichment relative to background abundances improves reliability of parameter estimates and reduces sensitivity to analytic error, as we have discussed previously (26, 37). As an empirical rule, any isotopic enrichment  $<0.0050$  (0.50 MPE) is problematic for MIDA calculations because of the large coefficient of variation that is unavoidable at such low enrichments (e.g., imprecision or inaccuracy of even  $\pm 0.0010$ – $0.0020$  represents a  $>20$ – $40\%$  coefficient of variation). A formal error analysis at  $0.0050$   $M_2$  enrichment in a sample polymer (palmitate) is shown (Fig. 7). These guidelines are dependent on the state of the technology, of course; if mass spectrometers improved accuracy and precision by an order of magnitude, the lower limit of acceptable incorporation would change accordingly.

The investigator can both estimate and influence the enrichments that are attained in an experiment (Fig. 8). The determinants of isotopomer enrichments in a polymer are  $n$  (the number of subunits),  $p$  (precursor pool labeling), and  $f$  (fractional synthesis). The value for  $f$  is generally one of the unknowns being investigated, so this cannot be manipulated by the investigator. The value for  $p$  can be controlled, however, by the dose of isotopic substrate administered. Here the balance is between avoiding an undesired substrate (non-tracer) effect in the pool of interest vs. achieving optimal enrichments in the product. It is sometimes possible to select the value of  $n$  that will be analyzed (e.g., for peptide fragments of a protein into which leucine or another labeled amino acid has been incorporated, Ref. 28), but this is more often not adjustable. The general relationship between  $p$ ,  $n$ , and enrich-

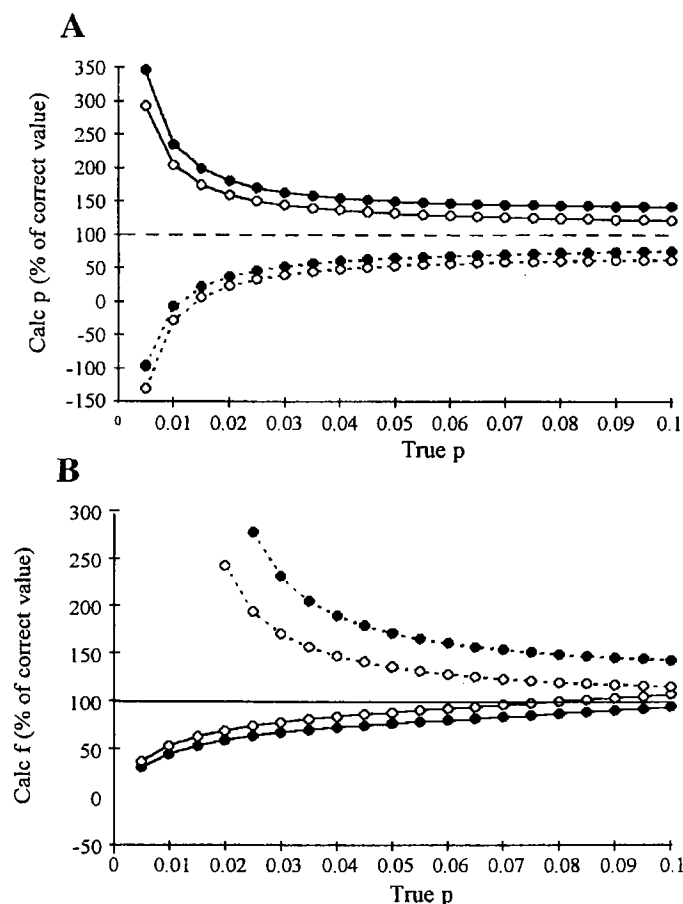


Fig. 7. Simulated error-analysis at low enrichment of an isotopomer in a polymer, in this case,  $M_2$  of methyl-palmitate at 0.0050 isotopic enrichment. Effects on calculated parameters ( $p$  and  $f$ ) of an analytic error of  $\pm 0.0010$  in  $\Delta A_2$  are simulated across the range of true  $p$  from 0.00 to 0.10 (at steps of 0.005). Calculated values of  $p$  and  $f$  (calc.  $p$ , calc.  $f$ ) are shown as a percentage of correct values of  $p$  and  $f$ . Different models of analytic error are compared: e.g., misidentification of  $M_2$  vs.  $M_1$  isotopomers (●-●); altered linearity of detector resulting in quantitative error in  $M_2$  abundances only (○-○). Solid lines, added to  $M_2$ ; dashed lines, subtracted from  $M_2$ .

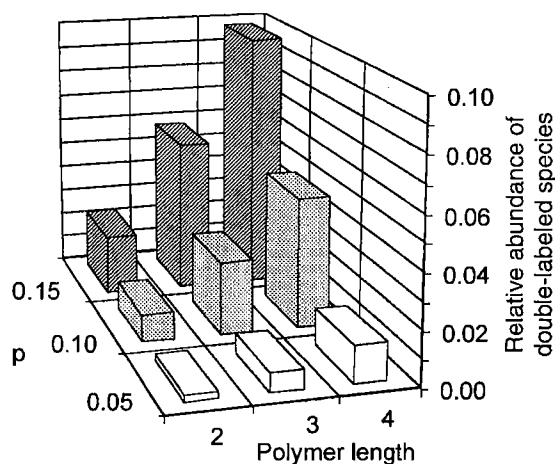


Fig. 8. Simulated relationship between  $n$  (polymer length as number of subunits),  $p$ , and relative abundance of double-labeled species in the case of a general polymer (for simplicity, natural abundance background is not included in the calculations).

ments of double-labeled species is shown in Fig. 8. Simulations (see APPENDIX) can provide useful information about the optimal dose of isotope to administer.

**Use of standard curves.** Most non-MIDA kinetic applications in stable isotope-mass spectrometry (e.g., dilution measurements of an infused tracer to determine metabolite flux) do not ultimately require instrument accuracy. Standard curves of the analyte can be made simply by mixing known amounts of labeled and unlabeled molecules in different proportions; the standard curve allows measured isotope enrichment to be converted to true proportions of labeled species present (tracer/tracee). The problem for biosynthetic measurements is that a simple standard curve cannot be easily made from available reagents, because there is no "standard" molecule for comparison. Each biological experiment results in a unique combination of  $p$  and  $f$ ; thus a unique pattern of mass isotopomers are present in the population of molecules. At a minimum, one would have to establish the performance of the instrument by creating a standard curve of  $p$  vs. measured ion abundances. It would be better to mix enriched molecules with unlabeled molecules to simulate end-product dilution ( $f$ ). Testing of three-dimensional standard curves in this manner requires substantial effort and analytic time, however, and does not correct for concentration effects (without imposition of another analytic dimension for the standards). Although mixtures of labeled standards can be useful for assessing instrument performance (3, 28 and previous discussion), the role of higher-dimensional standard curves in MIDA applications remains uncertain.

#### *Does the MIDA Calculation Algorithm Used Affect the Final Parameter Estimates, in Practice?*

As discussed above, the published algorithms for calculating biosynthetic parameters by MIDA from experimental data are similar in many respects and give essentially identical parameter estimates in the theoretical case (5). One potential difference is the capacity of different calculation approaches to adjust to instrument inaccuracy. Some calculation algorithms subtract measured baseline (natural abundance) data from enriched sample data before transforming sample data or fitting them to a model; other algorithms do not subtract measured baselines but correct for theoretical baselines or simply use the enriched sample data as they stand. In the former case, the goal is to account for instrument performance on a sample-to-sample basis; in the latter case, ideal instrument performance is assumed. The capacity of these two strategies to adapt to instrument inaccuracy can be simulated (Table 3). In this simulation, we considered several potential types of measurement inaccuracy: spillover between adjacent ions, constant contaminant at a particular mass-to-charge ratio ( $m/z$ ), nonlinearity of the detector at different  $m/z$  values. The impact on parameter estimates of using theoretical values vs. measured baselines was modeled. It is clear that use of measured natural abundance values in all cases reduces the error in parameter estimates, although not completely. This

Table 3. *Simulated analytic errors in glucose pentaacetate: effects of baseline values used on calculated values of  $f$*

Type of Error Involved	BSLN Used	Calculated $f$ (% actual value)
5% Spillover ( $M_1$ to $M_0$ )	Theoretical	80
	"Measured"	88
5% Spillover ( $M_0$ to $M_1$ )	Theoretical	126
	"Measured"	114
Constant contamination (0.0050 in $M_1$ )	Theoretical	107
	"Measured"	100

Simulations were carried out by imposing a unidirectional constant fraction error (5% of  $M_1$  into the  $M_0$  isotopomer or 5% of  $M_0$  into the  $M_1$  isotopomer) or a constant value error (0.0050 of the total envelope into the  $M_1$  isotopomer) on labeled glucose pentaacetate ( $p = 0.15$ ).  $\Delta A_1$ ,  $\Delta A_2$ , and parameter estimates  $p$  and  $f$  were then determined, after subtraction of either theoretical baseline (BSLN) values for glucose pentaacetate ions monitored [mass-to-charge ratio ( $m/z$ ) 331 to 333, see text] or "measured" BSLN values (i.e., imposing the same instrumental error as for labeled glucose).

conclusion was also reported by Chinkes et al. (5). Thus baseline corrections cannot substitute for accurate quantitation by mass spectrometers but may reduce error in the final calculated values.

#### *Is There Experimental Evidence Supporting or Contradicting the MIDA Model?*

A general measurement technique such as MIDA needs to be tested under controlled conditions both in vitro and in vivo. Strong experimental evidence in support of the method has been generated in vitro. Lee et al. (22) synthesized glucose-pentaacetate from [ $^{13}\text{C}$ ]acetic-anhydride at known enrichments and then imposed dilutions by known amounts of unlabeled glucose-pentaacetate. The calculated values of  $p$  and  $f$  were nearly identical to the expected values. We (28) performed an analogous test by synthesizing an oligopeptide (SVVLLLR) from [ $^2\text{H}_3$ ]leucine at known enrichments and imposing subsequent dilutions by unlabeled peptide. Again, observed values were extremely close to predicted values. The combinatorial probability-mass isotopomer quantification model is clearly an accurate theoretical description of polymerization biosynthesis under controlled conditions.

Testing a method definitively in vivo is not as straightforward. Comparisons with alternate methods, e.g., the close similarity of lipogenic estimates by  $^2\text{H}_2\text{O}$  and MIDA (reviewed in Ref. 16) or of cholesterologenesis by sterol balance and MIDA (25) are not definitive, because the comparison methods may be flawed as well. A useful result was presented by Kelleher et al. (20): cholesterologenesis approached the 100% value expected in exponentially growing cells in culture.

An important issue to discuss is the proper interpretation of apparently dissonant results with an in vivo method like MIDA. A recent example is the intriguing observation of Aarsland et al. (1) that total hepatic lipogenesis (based on MIDA on circulating very low density-triglyceride fatty acids) was less than 1/20th of the rate of net whole body lipogenesis (based on indirect calorimetry) during massive carbohydrate overfeeding

in humans. Because net lipogenesis (synthesis minus oxidation) cannot be greater than unidirectional lipogenesis, is this evidence that MIDA failed to give a physiologically possible answer? This conclusion would be incorrect: the authors (1) instead concluded that lipogenesis occurred in tissues or sites not immediately communicating with circulating very low density lipoprotein triglycerides during an 8-h isotope infusion. Thus lipogenesis in adipose tissue or lipogenesis entering the hepatic cytosolic storage pool could account for the unmeasured lipogenesis in the whole body. Physiological explanations must be excluded when the validity of MIDA is evaluated in vivo.

Analytic factors must also be excluded, particularly those related to instrument accuracy and concentration effects on measured isotope abundances (10, 11, 30). In our view (7, 26), analytic and experimental design factors may to some extent explain recent discordant results (32) of the use of MIDA for gluconeogenesis. Although this particular question has not been resolved, in general it is essential for manuscripts testing MIDA to demonstrate accuracy, avoid abundance-sensitivity effects, and explicitly describe the measures used to do so. Claims that a biosynthetic system is not adequately described by a combinatorial model need to be carefully evaluated using criteria other than just dissonant final estimates.

#### SOME FUTURE DIRECTIONS FOR MIDA

A number of future directions can be considered for combinatorial probability techniques. Applications with large, heteronuclear polymers (e.g., proteins, polynucleotides) represent a challenge but potentially include some of the most important molecules in biology. Special problems are introduced for heteronuclear polymers, such as proteins, for example, because their large size and composition (including 20 or more amino acid subunits) require the investigator to identify repeats of a particular homonuclear subunit and then isolate this subunit from the intact molecule (28). Although techniques such as electrospray ionization are able to introduce intact proteins into the gas phase for mass spectrometry, resolution of individual mass isotopomers is not generally possible with most mass analyzers, especially for the multiply charged species generated by electrospray ionization. We (28) have successfully used the strategy of enzymatically hydrolyzing proteins to proteolytic fragments, and then collecting a selected fragment that includes a homonuclear stretch (e.g., 3–5 leucines out of a 7- to 15-amino acid stretch). An analogous approach might be applicable to DNA or RNA samples, which have the advantage of containing only four different subunits.

Another potentially powerful direction for MIDA may be to exploit information about  $p$ , not just  $f$ . Establishing the timing of biosynthetic events (e.g., during embryonic development) might be possible by fingerprinting polymers according to their precursor pool enrichment after generating a time gradient within the precursor pool. Measuring the tissue location of a biosynthetic event (e.g., whether cholesterol in high-

density lipoprotein was synthesized in peripheral tissues or the liver, representing reverse or forward cholesterol transport) might also be possible by analyzing nonoverlapping values of  $p$  within an isotopomeric envelope, if label can be delivered at different rates to different tissues. Testing or correcting for labeling gradients within a tissue may also be possible by analyzing different portions of an isotopomeric envelope for divergent values of  $p$  (Fig. 5C).

### Conclusion

An algorithm (see APPENDIX) for calculating fractional abundances of mass isotopomers in complex mixtures of labeled and natural abundance polymers can be used to construct reference tables that allow inference of the biosynthetic parameters  $p$  and  $f$ . The consequences of deviations from various assumptions of the MIDA model and the practical limits of the technique can also be tested in this manner. MIDA appears to be robust in the face of a number of deviations from its central assumptions. The combinatorial-probability approach imposes minor or no constraints with regard to the isotope-labeled substrate that can be administered, the length of the polymer that can be studied, the presence of contaminating  $M_{-1}$  fragment ions, the existence of inhomogeneity or isotopic disequilibrium in the monomeric precursor pool, the possibility of variations in natural  $^{13}\text{C}$  abundances, or the existence of isotope discrimination. More problematic is analytic inaccuracy, which can be partially diagnosed by incorrect baseline isotopomer abundances and which is often exacerbated by concentration-sensitivity. Whether analytic imprecision or deviations from the assumptions of the simple MIDA model have an impact on biosynthetic parameters in a particular case can be tested by individualized simulation.

### APPENDIX: GENERAL CALCULATION ALGORITHM FOR MEASURING POLYMERIZATION BIOSYNTHESIS BY MIDA<sup>2</sup>

Our previous descriptions of the MIDA method (13, 14, 16, 17) provided the underlying theoretical basis but presented only an overview of the mathematical approach. The actual computer-based calculation algorithm that we have used for the past several years has not been described since our original presentation of the technique. We will describe a systematic and general calculation algorithm that can be used with a personal computer to generate expected fractional abundances of mass isotopomers for polymers containing labeled subunits; we will then show how these theoretical fractional abundances can be made into a "reference table" from which one can convert experimental data on mixtures of natural abundance and isotopically enriched polymers into biosynthetic parameters ( $p$  and  $f$ ); and finally, we will demonstrate how to perform a MIDA biosynthetic calculation from experimental data. How the algorithm can be used to evaluate limitations of the technique or deviations from the model,

or to correct for certain deviations from the model, is described in the text of this article.

### Calculation Algorithm

We first present a formal mathematical algorithm for calculating the fractional abundances of mass isotopomers resulting from mixing natural abundance molecules with molecules newly synthesized from a pool of labeled monomers characterized by the parameter  $p$ . A mixture of this type can be fully characterized by  $f$ , the fraction new, and  $p$ . This algorithm is presented in a step-wise fashion, beginning with the simplest calculation, a molecule synthesized from a single element containing isotopes with the same fractional abundances that occur in nature and not mixed with any other molecules. We then proceed to molecules containing more than one element with all isotopes at natural abundance; then to nonpolymeric molecules containing different elements, some of which are in groups whose isotope composition is not restricted to natural abundance but is variable; then to polymeric molecules containing combinations of repeating chemical units (monomers), wherein the monomers are either unlabeled (containing a natural abundance distribution of isotopes) or potentially labeled (containing an isotopically perturbed element group); and finally to mixtures of polymeric molecules, composed of both natural abundance polymers and potentially labeled polymers, the latter containing combinations of natural abundance and isotopically perturbed units. The last-named calculation addresses the condition generally present in a biological system, wherein polymers newly synthesized during the period of an isotope incorporation experiment are present along with preexisting natural abundance polymers, and the investigator is interested in determining the proportion of each that is present to infer synthesis rates or related parameters.

*Molecules containing only a single element at the fractional isotope abundances that occur in nature.* A first step is to calculate the isotope pattern of a hypothetical molecule composed of the element  $I$  with the fractional abundances  $^0I = 0.989$ ,  $^1I = 0.01$ ,  $^2I = 0.001$ . In the limit the molecule is a single atom, and the isotope pattern in a collection of element  $I$  is obtained directly from the fractional abundances of the isotopes:  $A_0 = 0.989$ ,  $A_1 = 0.01$ , and  $A_2 = 0.001$ . However, for a molecule composed of  $N$   $I$  atoms ( $I_N$ ), the mass isotopomer distribution (isotope pattern) is obtained from the multinomial distribution. This distribution is a function of two variables:  $N$  (the number of  $I$  atoms) and the fractional abundances of the isotopes, which in this case are not perturbed from natural abundances. The contribution at each mass unit is obtained by individually summing the fractional abundances of all the isotopologues with that particular nominal mass. In this fashion one obtains the fractional abundance that would be observed at that mass with a mass spectrometer operating at unit mass resolution, arising from the conglomerate of isotopologues with that nominal mass. The frequency ( $F$ ) of any given isotopologue in the summed total abundance of all of the isotopologues is given by the multinomial distribution [as adapted from *CRC Standard Mathematical Tables* (26th ed., pg 519); see Eq. A1]. For example, the probability of observing the  $I_8$  molecule composed of  $^0I_2^1I_5^2I_1$  is calculated as follows

$$F(^0I_2, ^1I_5, ^2I_1) = F(2,5,1) = \frac{8!}{2!5!1!} \times (.989)^2 (.01)^5 (.001)^1 = 1.64 \times 10^{-11}$$

<sup>2</sup> Richard A. Neese, Dennis Faix, Kenneth Caldwell, and Marc K. Hellerstein were the authors of the appendix. Dr. Caldwell is deceased as of August 1997.

Note that this is the fractional abundance for one particular isotopologue of nominal mass  $m_i$  due to the isotopologue  ${}^0I_2{}^1I_5{}^2I_1$ , not the fractional abundance of the particular mass isotopomer ( $M_i$ ) for the molecule, since there are other isotopologues that have a nominal mass of  $m_i$  (e.g.,  ${}^0I_3{}^1I_3{}^2I_2$ ). Thus the multinomial distribution allows the calculation of the probability or fractional abundance of any unique combination of isotopes (any isotopologue) in a molecule composed of a single element. To obtain the final probability of any particular mass isotopomer, one must sum the probabilities given by the multinomial distribution for each of the possible isotope combinations of the element. We note here that  $I_8$  would become an element group (defined as a group of atoms in a molecule attributed to a given element and sharing the same fractional abundances of the isotopes of the element; the fractional abundances of isotopes of the element, whether natural or perturbed, are by definition the same for all atoms in an element group, and a molecule may contain more than one element group for a given element). This element group is part of a molecule if that larger molecule contained eight  $I$  atoms and each of the  $I$  atoms were from the same isotopic pool.

*Molecules containing more than one element at natural isotope abundances.* Next, suppose that probability distribution calculations have been done for the element groups  $C_2$ ,  $H_4$ , and  $O_2$  as if they were isolated molecules, each composed of a single element, as described in the previous section. How then is the mass isotopomer distribution for  $C_2H_4O_2$  (acetic acid) obtained? One way is to multiply the fractional abundances of the components of each element group together and then to sum the fractional abundances of the specific isotopologues according to their nominal mass. Accordingly, to obtain the  $M_1$  components of acetic acid, the  $M_1$  component of each element group is combined with the  $M_0$  components of the others: ( ${}^{13}C{}^{12}C$ ,  ${}^1H_4$ ,  ${}^{16}O_2$ ), ( ${}^{12}C_2$ ,  ${}^1H_3{}^2H$ ,  ${}^{16}O_2$ ), and ( ${}^{12}C_2$ ,  ${}^1H_4$ ,  ${}^{16}O{}^{17}O$ ). Each of these unique combinations has a probability of occurring that can be calculated by using the multinomial distribution (Eq. A1); they are multiplied together to obtain the fractional abundances of the isotopologues (see Eq. A2). The fractional abundance ( $A_i$ ) of the  $M_1$  mass isotopomer of acetic acid is then the sum of these fractional abundances, representing the sum of all the isotopologue fractional abundances (Eq. A2d).

Although calculating the fractional abundances of individual isotopologues and summing them according to mass is a reasonable approach for obtaining a mass isotopomer distribution in simple molecules, a systematic, step-by-step algorithm is needed for more complex molecules. We now describe an algorithm, first in general terms and then by using acetic acid as an example. The strategy, alluded to in the previous example, is to disaggregate the molecule conceptually into element groups (e.g., C's, H's, O's), calculate the mass isotopomer distributions for each element group, and then multiply the appropriate components of each element group together in a systematic fashion to attain the final

mass isotopomer distribution of the molecule. A computer algorithm can do this by exhaustively combining all of the components of the first two element groups until the mass isotopomer distribution corresponding to the hypothetical molecule composed of those two element groups is obtained, and then by recording these abundances in an array. The process is then repeated for the next element group by combining its fractional abundances with those in the array. The resulting fractional abundances in the new array are then combined with the fractional abundances of the components of the next element group, and so on. A key aspect of this procedure is specifying all the combinations that contribute to the fractional abundances at each nominal mass during each iteration of the algorithm. For example, only  $M_0$  values from the array and the element group contribute to the  $M_0$  values in the new array, so they are multiplied together; however,  $M_1$  and  $M_0$  values contribute to the new  $M_1$ , so they must be multiplied each time. Likewise,  $M_0$ ,  $M_1$ , and  $M_2$  values from the two element groups contribute to the new  $M_2$  values, so  $M_0$  values are cross multiplied with  $M_2$  values, and  $M_1$  values are multiplied together. When all of the element groups have been combined together, the final distribution of fractional abundances for the whole molecule is attained. In the computer subroutine, this multiplication and summing process is carried out by use of loops, one nested inside the other. Because each addition of an element group increases the number of mass isotopomers and because the fractional abundances of mass isotopomers greater than approximately  $M_{10}$  are typically extremely low, an arbitrary limit may be set so that abundance calculations are not performed beyond this or some other set limit.

The example of acetic acid demonstrates how the computer algorithm is used (Table 4). The two element group distributions  $C_2$  and  $H_4$  are combined to obtain the mass isotopomer distribution of the intermediate  $C_2H_4$  molecule. This new distribution is in turn combined with the  $O_2$  distribution to form the final mass isotopomer distribution for acetic acid. The  $M_0$  value for  $C_2H_4$  (0.97771) is multiplied by the  $M_0$  value for  $O_2$  (0.99519) to yield the fractional abundance of the  $M_0$  mass isotopomer for acetic acid ( $A_0 = 0.97300$ ). Because only  $M_0$  mass isotopomers of the element groups contribute to that mass isotopomer for the final molecule, the calculations are complete for  $A_0$ , and the routine moves on to the next one, the  $M_1$  mass isotopomer. In this manner  $A_0$  to  $A_8$  are calculated (Table 4).

To facilitate combining two distributions, a single mass isotopomer distribution can be thought of as a set of values corresponding to the fractional abundances of each of the possible mass isotopomers of the molecule; any distribution (for an element group or molecule) is a set with  $n$  terms,  $M_0 \rightarrow M_{n-1}$ . Two distributions,  $A$  and  $B$ , will have  $n_A$  and  $n_B$  terms, respectively, whereas a distribution resulting from the combination of  $A$  and  $B$  will have  $n_{AB}$  terms, where  $n_{AB} = n_A + n_B - 1$ . The formula for combining two distributions,  $A$  and  $B$ , into a new single distribution  $AB$  is given in Eq. A3.

Table 4. Mass isotopomer abundance distribution calculations for natural abundance acetic acid ( $C_2H_4O_2$ )

Element Group	Mass Isotopomer ( $M_i$ )								
	$M_0$	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$	$M_7$	$M_8$
$C_2$	0.97832	0.02156	0.00012						
$H_4$	0.99938	0.00062	0.00000015	$1.5 \times 10^{-11}$	$5.9 \times 10^{-16}$				
$O_2$	0.995186	0.000738	0.00407	0.00000151	0.000000416				
$C_2H_4$	0.97771	0.02216	0.00013	$7.7 \times 10^{-8}$	$1.8 \times 10^{-11}$	$1.8 \times 10^{-15}$			
$C_2H_4O_2$	0.97300	0.02277	0.00413	0.00009	$4.6 \times 10^{-6}$	$9.3 \times 10^{-8}$	$5.5 \times 10^{-10}$	$3.2 \times 10^{-13}$	$7.4 \times 10^{-17}$



*Nonpolymeric molecules containing an element group in which isotope composition is perturbed from natural abundances and is variable.* Next, the calculation of mass isotopomer distributions is shown for nonpolymeric molecules that contain an element group enriched in a particular isotope, that is, perturbed from the natural fractional abundances of its isotopes. For a molecule  $I_xJ_y$ , where a number ( $w$ ) of the  $I$  element group is enriched in  $^{+1}I$ , the calculation is nearly the same as described already except that the  $w$  atoms of  $I$  must be considered as a separate element group ( $^*I$ ) from the other ( $x - w$ ) atoms of  $I$  in the molecule. The molecule then becomes  $^*I_wI_{x-w}J_y$ , with the appropriate distribution calculated for each element group and the element group distributions combined to generate the final distribution of fractional abundances of mass isotopomers in the whole molecule. The first step operationally, as described in the previous examples, is to break the molecule into element groups, calculating  $J_y$  and  $I_{x-w}$  from the natural isotope distributions of elements  $J$  and  $I$ . It is important to emphasize that  $J_y$  and  $I_{x-w}$  will have constant mass isotopomer abundance distributions even as the enriched elemental group ( $^*I$ ) changes. This invariant portion will be referred to here as the constant or invariant mass isotopomer abundance distribution of the molecule, and it can be treated as though it were a chemical derivatizing agent attached to the variable moiety. Dividing the isotopically perturbed molecule into constant and variable moieties greatly simplifies the calculation algorithm and is the motivation for dividing the elements in a molecule into element groups (one of which may be constant while the other is variable).

After the constant distribution is calculated, the distribution in the variable moiety is calculated. It is useful to divide the  $^*I_w$  element group itself into two populations. The first population is the proportion ( $p$ ) of the element group that is the perturbed portion, made up of isotopically labeled atoms (e.g.,  $^1I$ ). The second population is the proportion  $1 - p$  of the element group that is made up of the natural abundance atoms. The abundance distribution of the labeled proportion ( $^1I_w$ ) is simple<sup>3</sup>:  $A_w = 1.0$ . For the natural abundance proportion of the element group ( $^{na}I_w$ ), the distribution is calculated in the manner described above, by using the multinomial distribution (Eq. A1) with natural abundance values for  $w$   $I$  atoms. Because these two populations are physically mixed at the site of synthesis of the molecule from  $^*I_w$ , the combined distribution must be calculated. This combined distribution is calculated by correcting for the proportion of each present at a given value of  $p$ : the labeled distribution is multiplied by  $p$  and the natural abundance distribution by  $1 - p$ , and the fractional abundances of the corresponding nominal masses are summed. Thus a weighted distribution is generated for  $^*I_w$ . This new distribution is the fractional abundance distribution for  $^*I_w$  at that particular value of  $p$  and varies as a function of  $p$ . The  $^*I_w$  distributions over a broad range of values for  $p$  are then combined with the constant distribution to calculate the final molecular distribution as a function of  $p$ .

By way of illustration, the mass isotopomer fractional abundance distribution is calculated for a theoretical pool of  $[^2H_3]$ leucine molecules, which has been synthesized chemically. The element groups for these leucine molecules ( $C_6H_{13}O_2N$ ) are as follows:  $C_6$ ,  $H_{10}$ ,  $O_2$ , and  $N$  (the constant portion, all at natural isotope abundances) and  $^*H_3$  (the perturbed or variable portion, with a variable distribution).

The  $C_6$ ,  $H_{10}$ ,  $O_2$ , and  $N$  distributions are calculated individually and combined (Eq. A3) to form the constant distribution, which then must be combined with the distribution of the variable portion of the molecule. For the variable portion, the fractional abundance of  $^2H_3$  in  $^*H_3$  varied from zero (i.e., all  $^*H_3$  values are at natural abundance) up to a reasonable point of interest, when the fractional abundances of  $M_0 \rightarrow M_3$  are calculated for each value of  $p$ . For example, if  $p = 0.04$  (4% of the  $^*H_3$  element group consists of  $^2H_3$ ), 96% will have the natural abundance distribution for  $^*H_3$  ( $A_0 = 0.99953$ ,  $A_1 = 0.00047$ ,  $A_2 = 0.000$ ,  $A_3 = 0.000$ ), and 4% will have a distribution for  $^*H_3$  of  $A_3 = 1.0$ . Combining these two distributions by using a weighted average, or linear combinations, gives a final  $^*H_3$  distribution at  $p = 0.04$  of  $A_0 = 0.95955$ ,  $A_1 = 0.00045$ ,  $A_2 = 0.00$ , and  $A_3 = 0.04$ . This distribution is then combined with the constant distribution to obtain the final molecular isotopomer distribution, as seen in Table 5 for  $p = 0.04$ . Each  $^*H_3$  distribution at a given value of  $p$  is combined with the constant distribution to give a distribution of mass isotopomer abundances for the molecule leucine as a function of  $p$  (Table 5).

*Polymeric molecules containing repeating monomeric subunits, with some of the monomeric subunits composed exclusively of element groups at natural isotope abundances while other monomeric subunits contain an isotopically perturbed element group.* The next step is to calculate the fractional abundances of mass isotopomers in polymers containing combinations of monomers, which themselves are either unlabeled (containing only element groups at natural isotope abundances) or labeled (containing an isotopically perturbed element group). The calculation procedure is an extension of the principles described so far. A general discussion is presented first, followed by an example of a peptide containing repeated subunits of leucine, either at natural abundance or containing  $[^2H_3]$ leucine (see above).

Consider a polymer composed of  $I_xJ_y$  subunits, i.e.,  $(I_xJ_y)_z$ . If a number ( $w$ ) of  $I$  atoms ( $w < x$ ) in each  $I_xJ_y$  subunit are enriched in  $^{+1}I$  ( $^*I$ ), whereas all other atoms are at natural isotope abundances ( $I$  and  $J$ ), the calculation for each subunit  $I_xJ_y$  is as just described for a nonpolymeric molecule that contains an element group enriched in a particular isotope. The mass isotopomers of the polymer  $(I_xJ_y)_z$  then can be calculated and will depend upon  $p$ . Several mathematically interchangeable algorithms exist for calculating the distribution of mass isotopomer abundances for a polymer of this type. The calculation approach that we believe to be the simplest computationally is to treat the isotopically perturbed element group in the polymer as a single, discrete (inseparable) unit, with its own mass isotopomer abundance distribution that is then combined with the constant distribution of the remainder of the polymer. The first step, as with the described nonpolymeric leucine example, is to break the polymer into element groups. The  $I$  positions that can be enriched are treated as a separate polymer ( $^*I_w$ )<sub>z</sub>. This is convenient because the mass isotopomer abundance distribution of ( $^*I_w$ )<sub>z</sub> will vary as a function of  $p$ , whereas the remaining element groups  $J_{yz}$  and  $I_{(x-w)z}$  remain at natural isotope abundances and therefore have a constant abundance distribution of mass isotopomers, even as the percentage of isotopically perturbed subunits changes. This constant or invariant distribution of the polymer is treated identically as in simple, nonpolymeric molecules (Table 4). Therefore, one can first calculate the mass isotopomer abundance distributions for the element groups that are invariant. These are then combined, one at a time, until the constant distribution is calculated.

<sup>3</sup> We will assume 100% isotopic labeling by  $^{+1}I$  in the material. If the material is  $<100\%$   $^{+1}I$ , the actual distribution of  $^1I_w$  is calculated by using Eq. A3.



Table 5. Mass isotopomer fractional abundance distributions for leucine pool at varying proportions ( $p$ ) of labeled [ $^2\text{H}_3$ ]leucine

Mass Isotopomer	Fractional Abundance						
	$M_0$	$M_1$	$M_2$	$M_3$	$M_4$	$M_5$	$M_6$
Nominal mass ( $m/z$ )	131	132	133	134	135	136	137
Exact mass	131.10	132.10	133.10	134.10	135.10	136.10	137.10
$p$							
0.00	0.92657	0.06723	0.00588	0.00031	0.00001	0.00000	0.00000
0.02	0.90804	0.06588	0.00577	0.01884	0.00135	0.00012	0.00000
0.04	0.88950	0.06454	0.00565	0.03738	0.00269	0.00024	0.00001
0.06	0.87097	0.06319	0.00553	0.05591	0.00402	0.00035	0.00002
0.08	0.85244	0.06185	0.00541	0.07444	0.00536	0.00047	0.00002
0.10	0.83391	0.06050	0.00530	0.09298	0.00669	0.00059	0.00003

Calculations for the variable part of the polymer ( $*I_w$ ) $_z$  are more complicated. Its fractional abundance distribution will depend on the number of labeled subunits ( $*I_w$ ) that happen to be incorporated. If we call the number of labeled subunits  $\sigma$ , this element group may be composed entirely of labeled subunits ( $\sigma = z$ ), or it may have 1, 2, 3 and so on, labeled subunits ( $\sigma = 1, 2, 3, \dots$ ), or it may have no labeled subunits ( $\sigma = 0$ ). The remaining unlabeled subunits ( $z - \sigma$ ) contain only elements at natural isotope abundances. The general formula for this variable element group is therefore ( $*I_w$ ) $_{\sigma} (I_w)_{z-\sigma}$ .

Accordingly, the next step is to calculate all the possible mass isotopomer distributions for  $\sigma = 0$  to  $z$ . If  $\sigma = 0$ , then the calculation is simply that of inserting the natural abundance values into the multinomial distribution to obtain a distribution for  $wz$   $I$  atoms ( $I_{wz}$ ), as was calculated for the constant part of the polymer. If  $\sigma = 1$ , then the distribution is obtained by calculating the natural abundance distribution for  $w(z-1)$   $I$  atoms [ $I_{w(z-1)}$ ] and the distribution for  $*I_w$  ( $A_1 = 1.0$ ) and combining these two distributions (Eq. E3) to generate the distribution of the variable portion of a polymer with a single labeled subunit. If  $\sigma = 2$ , the natural abundance distribution [ $I_{w(z-2)}$ ] is combined with the  $*I_w$  distribution ( $A_2 = 1.0$ ) in the same manner. The process is repeated up to  $\sigma = z$ . If all the subunits are labeled ( $\sigma = z$ ), then the distribution for  $*I_{wz}$  becomes  $A_z = 1.0$ , and all other mass isotopomer fractional abundances are zero. Stated in general terms, when the isotopomer distribution for an element group containing  $\sigma$  perturbed or labeled subunits is calculated, the distributions for  $*I_{w\sigma}$  (labeled) and  $I_{w(z-\sigma)}$  (natural abundance) are calculated and combined to give the distribution for that subunit combination in the element group.

These distributions for each value of  $\sigma$  serve as reference distributions and need only to be calculated once for any given polymeric molecule containing a particular labeled element group. They do not, however, by themselves reveal the actual distribution in the polymeric molecule, because the proportion of each of these theoretical species ( $\sigma = 0, 1, 2, 3, \dots$ ) actually present must also be calculated, and this is determined by  $p$ , the fractional abundance of isotopically labeled subunits in the precursor pool. The multinomial, or, in this example, the binomial expansion (Eq. A1), is used to calculate the distribution of subunits in the element group as a function of  $p$  and  $z$ , the total number of subunits in the element group (Eq. A4). Each value of  $p$  will be associated with unique proportions of  $\sigma = 0, 1, 2, 3, \dots$  subunits in the element group (Table 6A).

Accordingly, one then multiplies the frequency of each subunit combination (i.e., the coefficient) by the mass isotopomer abundances characteristic of that combination (i.e., the distribution) to generate a final, weighted distribution. This is the ( $*I_w$ ) $_z$  mass isotopomer abundance distribution for

that value of  $p$  (Table 6B). This distribution is combined with the distribution of the invariant moiety of the polymer (using Eq. A3) to give a molecular distribution for that value of  $p$ . The process is repeated for every  $p$  of interest ( $0 \rightarrow .xxx$  at steps .00x) until a table of mass isotopomer fractional abundances of the polymer as a function of  $p$  is obtained (Table 7). During iterations as  $p$  is varied, only the probability of each subunit combination needs to be recalculated before multiplying, summing, and combining to form the fractional abundance distribution of mass isotopomers of the molecule at that value of  $p$  (Table 6B). The reference table obtained (Table 7) reveals data for  $p$  vs. isotopomer abundances in the intact SVVLLLR peptide. It should be emphasized that each molecule analyzed by MIDA requires its own reference table. This needs to be calculated only once, however, across the range of values for  $p$ ; the table can then be used for all subsequent measurements on this molecule. Computational intensity is substantial but applies most often at the beginning of work on a biosynthetic problem.

**Mixtures of polymers.** Finally, one can consider mixtures of polymers, as might be present during a physiological isotope incorporation experiment in vivo, wherein newly synthesized polymers are being added to a preexisting pool of polymers as part of the process of turnover. This calculation consists simply of generating linear combinations of the fractional abundances of each mass isotopomer.<sup>4</sup> Thus one multiplies the fractional abundances of mass isotopomers in each polymer population times the proportion of the polymer pool represented by each population, and one sums the results. In this manner, the fractional abundances of each mass isotopomer in the mixture are weighted for the proportion of each polymer present. The net result is a linear combination of mass isotopomer abundances (Eq. A5).

#### Generation of Reference Tables for Inferring Biosynthetic Parameters from Experimental Data

**Rationale for use of  $\Delta$ fractional abundances in reference tables.** As discussed in the text of this article and previously (13, 15, 20, 21), inference of the fractional biosynthetic contribution to a polymeric pool ( $f$ ) requires knowledge of the proportion ( $p$ ) of precursor subunits that were isotopically labeled after introduction of a labeled monomer. The mathematical reason why  $p$  must be known to establish  $f$  is evident by graphical analysis (see accompanying manuscript, Fig. 1C). The universe of possible fractional abundances for any mass isotopomer in a polymer can be seen to depend on the

<sup>4</sup> This holds unless an incomplete ion spectrum is being monitored, in which case a correction must be applied to restore the molar combinations to linearity (see text and Eq. A9b).

Table 6. Mass isotopomer distributions in peptide SVLLLLR (singly charged, protonated) due to incorporation of [ $^2\text{H}_3$ ]leucine at  $p = 0.10$ 

A. Combinations of natural abundance leucine and [ <sup>2</sup> H <sub>3</sub> ]leucine subunits at p = 0.10 (coefficients)											
[ <sup>2</sup> H <sub>3</sub> ]leucine Units (σ)	Atomic Elemental Formula			Monoisotopic Mass, μm			Fractional Contribution to Peptide (Coefficient)				
0	C <sub>37</sub> N <sub>10</sub> O <sub>9</sub> H <sub>71</sub> ( <sup>2</sup> H <sub>0</sub> )			799.6			$\left(\frac{3!}{3!0!}\right) \times (0.9)^3(0.1)^0 = 0.729$				
1	C <sub>37</sub> N <sub>10</sub> O <sub>9</sub> H <sub>68</sub> ( <sup>2</sup> H <sub>3</sub> )			802.6			$\left(\frac{3!}{2!1!}\right) \times (0.9)^2(0.1)^1 = 0.243$				
2	C <sub>37</sub> N <sub>10</sub> O <sub>9</sub> H <sub>65</sub> ( <sup>2</sup> H <sub>6</sub> )			805.6			$\left(\frac{3!}{1!2!}\right) \times (0.9)^1(0.1)^2 = 0.027$				
3	C <sub>37</sub> N <sub>10</sub> O <sub>9</sub> H <sub>62</sub> ( <sup>2</sup> H <sub>9</sub> )			808.6			$\left(\frac{3!}{0!3!}\right) \times (0.9)^0(0.1)^3 = 0.001$				
B. Fractional abundances of mass isotopomers M <sub>0</sub> to M <sub>10</sub> (m/z 799 to 809) in peptide at p = 0.10 for [ <sup>2</sup> H <sub>3</sub> ]leucine											
[ <sup>2</sup> H <sub>3</sub> ]leucine units, σ	M <sub>0</sub> (799)	M <sub>1</sub> (800)	M <sub>2</sub> (801)	M <sub>3</sub> (802)	M <sub>4</sub> (803)	M <sub>5</sub> (804)	M <sub>6</sub> (805)	M <sub>7</sub> (806)	M <sub>8</sub> (807)	M <sub>9</sub> (808)	M <sub>10</sub> (809)
0											
Distribution	0.621963	0.285353	0.075466	0.014608	0.002273	0.000299	0.000034	0.000003	0.000000	0.000000	0.000000
Coefficient	0.729	0.729	0.729	0.729	0.729	0.729	0.729	0.729	0.729	0.729	0.729
Subtotal	0.453411	0.208022	0.055015	0.010649	0.001657	0.0000218	0.000025	0.000002	0.000000	0.000000	0.000000
1											
Distribution				0.622255	0.285195	0.075368	0.014579	0.002267	0.000298	0.000034	0.000003
Coefficient				0.243	0.243	0.243	0.243	0.243	0.243	0.243	0.243
Subtotal	0.000000	0.000000	0.000000	0.151208	0.069302	0.018314	0.003543	0.000551	0.000072	0.000008	8.4 × 10 <sup>-7</sup>
2											
Distribution							0.622546	0.285037	0.075270	0.014551	0.002261
Coefficient							0.027	0.027	0.027	0.027	0.027
Subtotal	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.016809	0.007696	0.002032	0.000393	0.000061
3											
Distribution										0.622837	0.284879
Coefficient										0.001	0.001
Subtotal	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.001024	0.000285
Totals											
Fractional abundance	0.453411	0.208022	0.055015	0.161857	0.070959	0.018532	0.020377	0.008249	0.002104	0.001024	0.000347
%Relative abundance	100.00	45.88	12.13	35.70	15.65	4.09	4.49	1.82	0.46	0.23	0.08

Distribution of the invariant moiety of the molecule is corrected for the proportion of each combination of labeled to unlabeled leucine subunits (the coefficient), to generate abundances of each mass isotopomer.

two variables,  $p$  and  $f$ . By envisioning a plane drawn at the fractional abundance value for any  $A_x$  (i.e., 0.15, Fig. 1C), it is apparent that an infinite number of solution pairs of  $p$  and  $f$  could result in any given value of  $A_x$ . In a nonlinear system in which both  $p$  and  $f$  are unknown and variable, such as an in vivo biosynthetic system, it is therefore necessary to solve for  $p$  to linearize the relationship between fractional abundance  $A_x$  and  $f$ . It is then possible to solve for  $f$  by simple algebraic means.

The calculation problem that has to be solved for mixtures is that the polymers at natural isotope abundances also contribute to the abundances of mass isotopomers. There is, however, a simple way of correcting for natural abundance contributions and for linearizing the relationship between  $f$  and  $A_x$  (Eqs. A5–A7): when expressed as  $\Delta$ fractional abundances, ratios among mass isotopomers are constant and independent of  $f$  (Eq. A7c). It follows that the isotopomer pattern of the population of newly synthesized molecules can be established by expressing the data as ratios among  $\Delta$ fractional abundances. The value for  $p$  can then be inferred directly from reference tables (see Table 7), and the value for  $f$  is calculated easily by application of the precursor-product relationship. The calculation advantage of subtracting mea-

sured baselines has also been recognized and incorporated into the algorithms of Lee et al. (22) and Chinkes et al. (5), although other published algorithms do not include a baseline correction (21).

Another mathematical point concerning the use of  $\Delta$ fractional abundances is worth mentioning. The sum of depletions (negative values) plus enrichments (positive values) compared with natural fractional abundances for all the mass isotopomers analyzed must add up to zero, because fractional abundances by definition add up to 1.0 in the mass isotopomer envelope analyzed, whatever the value for  $p$  or  $f$ . The absolute value of the enrichments ( $+\Delta A_x$ , or  $EM_x$ ) plus depletions ( $-\Delta A_x$ , or  $DM_x$ ), or the degree of perturbation from natural abundance values ( $\Delta A_x^\circ$ ), increases as  $p$  or  $f$  increases, but  $EM_x$  is always cancelled out by  $DM_x$ .

#### Conversion of Experimental Data into Biosynthetic Parameters

The sequence followed in converting experimental data into biosynthetic parameters is as follows. The ratio (R) of  $\Delta$ fractional abundances of mass isotopomers in a mixture of polymers is used to calculate  $p$  and  $\Delta A_x^\circ$  by using the

Table 7. Theoretical abundances of selected mass isotopomers in peptide SVVLLLR as a function of  $p$ 

$p$	$M_0$	$M_3$	$M_6$	$\Delta A_3^{\infty}$	$\Delta A_6^{\infty}$	$\Delta A_6^{\infty}/\Delta A_3^{\infty}$
0.00	0.97702	0.02293	0.00005			
0.01	0.94800	0.05099	0.00102	0.02806	0.00096	0.03427
0.02	0.91957	0.07791	0.00252	0.05498	0.00247	0.04485
0.03	0.89175	0.10371	0.00455	0.08078	0.00449	0.05564
0.04	0.86451	0.12840	0.00709	0.10547	0.00703	0.06667
0.05	0.83787	0.15202	0.01011	0.12909	0.01006	0.07793
0.06	0.81181	0.17458	0.01362	0.15165	0.01356	0.08944
0.07	0.78633	0.19609	0.01758	0.17317	0.01753	0.10121
0.08	0.76142	0.21659	0.02198	0.19367	0.02193	0.11323
0.09	0.73709	0.23610	0.02681	0.21317	0.02676	0.12553
0.10	0.71332	0.25463	0.03205	0.23170	0.03200	0.13811
0.11	0.69011	0.27220	0.03769	0.24927	0.03763	0.15097
0.12	0.66746	0.28884	0.04370	0.26591	0.04365	0.16414
0.13	0.64535	0.30457	0.05008	0.28165	0.05003	0.17762
0.14	0.62378	0.31942	0.05681	0.29649	0.05675	0.19142
0.15	0.60274	0.33339	0.06387	0.31046	0.06382	0.20556
0.16	0.58223	0.34652	0.07126	0.32359	0.07120	0.22004
0.17	0.56223	0.35882	0.07895	0.33589	0.07889	0.23488
0.18	0.54275	0.37032	0.08693	0.34739	0.08688	0.25009
0.19	0.52376	0.38104	0.09520	0.35811	0.09514	0.26569
0.20	0.50528	0.39099	0.10373	0.36806	0.10368	0.28169
0.21	0.48728	0.40020	0.11252	0.37728	0.11247	0.29810
0.22	0.46975	0.40869	0.12155	0.38577	0.12150	0.31495
0.23	0.45270	0.41648	0.13081	0.39355	0.13076	0.33225
0.24	0.43611	0.42359	0.14030	0.40066	0.14024	0.35003
0.25	0.41998	0.43003	0.14998	0.40711	0.14993	0.36829

Best-fit equation for calculating  $p$ :  $p = -0.0239 + 1.0236(\Delta A_6^{\infty}/\Delta A_3^{\infty}) - 1.0414(\Delta A_6^{\infty}/\Delta A_3^{\infty})^2 + 1.0160(\Delta A_6^{\infty}/\Delta A_3^{\infty})^3 - 0.8232(\Delta A_6^{\infty}/\Delta A_3^{\infty})^4 + 0.3783(\Delta A_6^{\infty}/\Delta A_3^{\infty})^5$ . Best-fit equation for calculating  $\Delta A_3^{\infty}$  as a function of  $p$ :  $\Delta A_3^{\infty} = 1.8234(p) - 3.694(p^2) + 2.0276(p^3) + 1.0549(p^4) - 1.7611(p^5)$ .

appropriate reference table (e.g., Table 7).  $\Delta A_x^{\infty}$  represents the  $\Delta A_x(\text{mixture})$  for 100% newly synthesized molecules at that value of  $p$ , i.e., the asymptotic  $\Delta A_x(\text{mixture})$  that could be attained at that value of  $p$ . The observed  $\Delta$ fractional abundance of any or all convenient mass isotopomers in the mixture [ $\Delta A_x(\text{mixture})$ ] is then compared with the maximal or asymptotic value,  $\Delta A_x^{\infty}$ , to generate  $f$ . It should be noted that both  $p$  and  $\Delta A_x^{\infty}$  are directly inferred from R in the reference table. That is, in practice one does not interpose an estimate of  $p$  to calculate  $\Delta A_x^{\infty}$  from experimental data, but  $\Delta A_x^{\infty}$  is derived directly from R in the reference table. If an incomplete ion spectrum is monitored (as, in this example,  $M_0$ ,  $M_3$ ,  $M_6$ ), the correction factor equation is applied for calculating  $f$  (Eq. A9b).

#### Sample Calculations

A sample of experimental data will now be presented. The example described is of an oligopeptide fragment of human serum albumin (SVVLLLR, or serine-valine-leucine-leucine-leucine-arginine), which can be produced by partial proteolytic degradation of human serum albumin by treatment with the enzymes trypsin and chymotrypsin (28). This oligopeptide contains three repeating leucine subunits and was synthesized in vitro from mixtures of natural abundance leucine and [ $^2\text{H}_3$ ]leucine subunits. The reference table for this compound, establishing the relationship between  $p$  and  $\Delta$ fractional abundances, must first be generated. The process is as described above. Briefly, the chemical composition of the molecule is established ( $\text{C}_{37}\text{N}_{10}\text{O}_9\text{H}_{71}$ ). The monoisotopic mass of this molecule is 799.6. The molecule is then divided into a constant moiety ( $\text{C}_{37}\text{N}_{10}\text{O}_9\text{H}_{62}$ ) and a variable moiety ( $^*\text{H}_9$ ). The fractional abundance distribution of mass isotopomers

for the constant moiety is calculated (all elements at natural abundance), and the fractional abundance distribution of mass isotopomers for each combination of labeled and natural abundance subunits in the variable moiety ( $\text{H}_9^2\text{H}_0$ ,  $\text{H}_6^2\text{H}_3$ ,  $\text{H}_3^2\text{H}_6$ ,  $\text{H}_0^2\text{H}_9$ ) is calculated (Table 6A). Then the fractional contribution to the peptide of each combination of labeled and unlabeled subunits is calculated at a given value of  $p$  (Table 6B). For  $p = 0.10$ , the likelihood of  $\text{H}_9^2\text{H}_0$  (zero units of labeled  $^2\text{H}_3 + 3$  units of natural abundance labeled  $\text{H}_3$ ) = 0.729, representing the proportion of  $M_0$  in the variable moiety; the likelihood of  $\text{H}_6^2\text{H}_3$  (1 labeled  $^2\text{H}_3$  unit + 2 natural abundance labeled  $\text{H}_3$  units) = 0.243, representing the proportion of  $M_{12}$  in the variable moiety, and so on. Each of these coefficients is then multiplied times the mass isotopomer fractional abundances in the corresponding combination of labeled and natural abundance  $\text{H}_3$  subunits and incorporated into a table of the final abundances of each mass isotopomer in the oligopeptide at  $p = 0.10$  (Table 6B). The same process may be iterated for  $p = 0.00$  to  $p = 0.20$ , at stepwise intervals of 0.001, for example (Table 7). The natural abundance values are then subtracted for each mass isotopomer at each value of  $p$  to generate  $\Delta$ fractional abundances ( $\Delta A_x^{\infty}$  or  $EM_x$ , Table 7). The result is a reference table for  $\Delta$ fractional abundances of mass isotopomers of SVVLLLR as a function of  $p$  ( $[^2\text{H}_3]$ leucine proportion in the precursor pool).

Experimental data from a mixture of isotopically perturbed and natural abundance SVVLLLR molecules can then be analyzed. The peak areas are converted to fractional abundances, assuming measurements are made on the complete ion spectrum ( $M_0 - M_6$ ). Baseline (natural abundance) experimental data are subtracted to generate  $\Delta$ fractional abundances. If, for example, the value of  $\Delta A_3$  (mixture) = 0.0941 and  $\Delta A_6$  (mixture) = 0.0214, then the ratio  $\Delta A_6(\text{mixture})/\Delta A_3(\text{mixture})$  was 0.227. The reference table equation (Table 7) indicates that  $p = 0.165$  and  $\Delta A_3^{\infty} = 0.2091$ . Next,  $f$  is calculated as  $\Delta A_3(\text{mixture})/\Delta A_3^{\infty} = 0.0941/0.2091$ , or 45%. In a biological experiment, the half-time ( $t_{1/2}$ ) of albumin could then be calculated (i.e.,  $f = 1 - e^{-kt}$ , and  $t_{1/2} = 0.69/k_3$ ).

#### Appendix Equations

##### Equation A1

$$F(^0\text{I}_{x_0}, ^1\text{I}_{x_1}, ^2\text{I}_{x_2}, \dots, ^{n-1}\text{I}_{x_{n-1}}) = F(x_0, x_1, x_2, \dots, x_{n-1}) = \frac{M}{\prod_{i=0}^{n-1} x_i!} \prod_{i=0}^{n-1} A_i^{x_i}$$

where  $x_i$  is the number of atoms of the  $^i\text{I}$  isotope,  $N$  is the total number of atoms ( $N = \sum_{i=0}^{n-1} x_i$ ),  $n$  is the number of isotopes of  $I$ , in this case three ( $^0\text{I}$ ,  $^1\text{I}$ , and  $^2\text{I}$ ), and  $A_i$  is the fractional abundance of  $^i\text{I}$ , where  $A_i > 0$  for  $i = 0, 1, 2, \dots, n-1$  and  $\sum_{i=0}^{n-1} A_i = 1$ .  $F$  is the probability (or statistical frequency) of observing the combination of events indicated. Taking acetic acid as an example

##### Equation A2a

$$\begin{aligned} F(^{13}\text{C}^{12}\text{C}, ^1\text{H}_4, ^{16}\text{O}_2) &= [p(^{12}\text{C}^{13}\text{C}) \times p(^1\text{H}_4) \times p(^{16}\text{O}_2)] \\ &= [p(M_1, \text{C}_2) \times p(M_0, \text{H}_4) \times p(M_0, \text{O}_2)] \\ &= (0.02156)(0.99938)(0.99519) \\ &= 0.021443 \end{aligned}$$

Equation A2b

$$\begin{aligned}
 F(^{12}\text{C}_2, ^1\text{H}_3^2\text{H}, ^{16}\text{O}_2) &= [p(^{12}\text{C}_2) \times p(^1\text{H}_3^2\text{H}) \times p(^{16}\text{O}_2)] \\
 &= [p(M_0, \text{C}_2) \times p(M_1, \text{H}_4) \times p(M_0, \text{O}_2)] \\
 &= (0.97832)(0.00062)(0.99519) \\
 &= 0.000604
 \end{aligned}$$

Equation A2c

$$\begin{aligned}
 F(^{12}\text{C}_2, ^1\text{H}_4, ^{16}\text{O}^{17}\text{O}) &= [p(^{12}\text{C}_2) \times p(^1\text{H}_4) \times p(^{16}\text{O}^{17}\text{O})] \\
 &= [p(M_0, \text{C}_2) \times p(M_0, \text{H}_4) \times p(M_1, \text{O}_2)] \\
 &= (0.97832)(0.99938)(0.00074) \\
 &= (0.000724)
 \end{aligned}$$

where  $p$  refers to the probability of that particular element group combination occurring, and where  $p(M_x, I_i)$  refers to the probability of  $M_x$  in the element group that follows.

Equation A2d

$$A_1 = 0.021443 + 0.000604 + 0.000724 = 0.02277$$

A more systematic approach for calculating mass isotopomer distributions is by disaggregating the molecule into its component element groups (e.g., C's, H's, O's), calculating the mass isotopomer distributions for each group, and then multiplying the appropriate components of each group together, systematically, to generate the mass isotopomer distribution of the molecule

Equation A3

$$\bigcup_{n=0}^{n_A + n_B - 1} p(M_n, AB) = \sum_{i=0}^n p(M_{n_A}, A) \times p(M_{n_B}, B)$$

$A$  and  $B$  can be either element group or molecular distributions, and the new, combined distribution is given by the set of mass isotopomer fractional abundances of  $M_0$  through  $M_{(n_A + n_B - 1)}$ . In this way an entire molecular distribution can be systematically calculated using Eq. A3 and element group mass isotopomer distributions.

The abundance distribution of labeled and unlabeled subunits that combine to form a polymer is calculated using the multinomial expansion

Equation A4

Abundance distribution of subunits in a polymer

$$= d(z, \sigma, p) = \sum_{\sigma=0}^z \frac{z!}{(z-\sigma)! \sigma!} (1-p)^{(z-\sigma)} p^\sigma$$

where  $\sigma$  is the number of labeled subunits present in the variable moiety of the polymer,  $z$  is the maximum number of subunits that can be labeled in the variable moiety of the polymer, and  $p$  is the fraction of isotopically labeled subunits in the subunit precursor pool.

An important mathematical feature of simple mixtures of molecular populations is that they combine linearly

Equation A5a

$$A_1 \text{ in mixture} = f(A_{1a}) + (1-f)(A_{1b})$$

Equation A5b

$$A_2 \text{ in mixture} = f(A_{2a}) + (1-f)(A_{2b})$$

where  $f$  is the fraction of polymer population  $a$  present (e.g., newly synthesized polymers),  $(1-f)$  is the fraction of polymer population  $b$  present (e.g., preexisting, unlabeled polymers), and  $A_{1a}$  and  $A_{1b}$  represent the fractional abundances of mass isotopomer  $M_x$  in populations  $a$  and  $b$ , respectively.

In a mixture of natural abundance and isotopically perturbed polymers, the fractional abundance of a mass isotopomer  $A_x$  is therefore (the general case in Eqs. A5a and A5b above)

Equation A5c

$$A_x (\text{mixture}) = f(A_{x_N} + \Delta A_x^\infty) + (1-f)(A_{x_N}) = A_{x_N} + f \cdot \Delta A_x^\infty$$

where  $A_{x_N}$  is the natural fractional abundance of the mass isotopomer  $M_x$ , and  $\Delta A_x^\infty$  is the change in fractional abundance of the mass isotopomer  $M_x$  in the isotopically perturbed polymers. We have elsewhere used the term  $EM_x$  in place of  $\Delta A_x^\infty$  or  $\Delta A_x$  (13, 14, 16, 17); the terms are interchangeable when  $\Delta A_x^\infty$  is a positive number. It should be noted that  $\Delta A_x^\infty$  and  $\Delta A_x$  may be positive or negative numbers.

Equation A5c restates the dependence of  $A_x$  on both  $p$  (since  $p$  determines  $\Delta A_x^\infty$  in the isotopically perturbed polymers) and  $f$ , in addition to explicitly recognizing the contribution from natural abundance molecules ( $A_{x_N}$ ) to the fractional abundance of the mixture.

Similarly, the fractional abundance of another mass isotopomer,  $A_y$ , in a mixture is

Equation A5d

$$A_y (\text{mixture}) = f(A_{y_N} + \Delta A_y^\infty) + (1-f)(A_{y_N}) = A_{y_N} + f \cdot \Delta A_y^\infty$$

where  $A_{y_N}$  is the natural fractional abundance of the mass isotopomer  $M_y$ , and  $\Delta A_y^\infty$  is the change in fractional abundance of  $M_y$  in isotopically perturbed polymers.

Because of the contributions from natural isotope abundances ( $A_{x_N}$  and  $A_{y_N}$ ), the mathematical relationship between  $A_x$  and  $A_y$  in the mixture (the internal pattern or distribution) is not uniquely determined by  $p$ , which determines the mass isotopomer pattern in the isotopically perturbed molecules, but is also determined by  $f$

Equation A6

$$A_x (\text{mixture})/A_y (\text{mixture}) = (A_{x_N} + f \cdot \Delta A_{x_N}^\infty)/(A_{y_N} + f \cdot \Delta A_{y_N}^\infty)$$

Accordingly, there are two unknowns that influence the internal relationships among mass isotopomers in a mixture of labeled and natural abundance polymers. More than one equally valid mathematical approach has been proposed to solve for  $p$  and  $f$  from experimental data in the presence of natural abundance contributions (6, 12, 14–17, 20, 21, 32). Our approach has been to linearize the relationship between  $p$  and  $A_x(\text{mixture})$  by expressing the results as  $\Delta$ fractional abundances (excesses or depletions compared with natural fractional abundance values). When  $\Delta$ fractional abundances are used, the internal relationship among mass isotopomers in the mixture [ $\Delta A_x(\text{mixture})/\Delta A_y(\text{mixture})$ ], when the changes in abundance are in the positive direction, becomes constant for  $p$  and independent of  $f$ , so that the initial reference table can be used regardless of the particular mixture present

Equation A7a

$$\Delta A_x (\text{mixture}) = [A_{x_N} + f \cdot \Delta A_{x_N}^\infty] - [A_{x_N}] = f \cdot \Delta A_{x_N}^\infty$$

Equation A7b

$$\Delta A_y (\text{mixture}) = [A_{y_N} + f \cdot \Delta A_{y_N}^\infty] - [A_{y_N}] = f \cdot \Delta A_{y_N}^\infty$$

## Equation A7c

$$\Delta A_x (\text{mixture}) / \Delta A_y (\text{mixture}) = f \cdot \Delta A_x^{\infty} / f \cdot \Delta A_y^{\infty} = \Delta A_x^{\infty} / \Delta A_y^{\infty}$$

Because  $\Delta A_x^{\infty}$  and  $\Delta A_y^{\infty}$  are defined as the change in fractional abundance in the newly synthesized or isotopically perturbed polymers only, Eq. A7c is independent of the dilution by unlabeled polymers, and thus only depends on the value of  $p$  that was present. The value of  $p$  is then calculated from the best-fit equation of  $p$  vs.  $\Delta A_x^{\infty} / \Delta A_y^{\infty}$  in the appropriate reference table. Using the example of the peptide SVVLLLR, if  $\Delta A_6 / \Delta A_3 = 0.19142$  (Table 7)

## Equation A8

$$p(\text{SVVLLLR}) = -0.0239 + 1.0236(0.19142) - 1.0414(0.19142)^2 + 1.0160(0.19142)^3 - 0.8232(0.19142)^4 + 0.3783(0.19142)^5 = 0.1400$$

The value of  $f$  is then calculated using  $\Delta A_x^{\infty}$ , at the value of  $p$  calculated from the  $\Delta A_x / \Delta A_y$  ratio present

## Equation A9a

$$f = \Delta A_x (\text{mixture}) / \Delta A_x^{\infty}$$

where  $\Delta A_x^{\infty}$  represents the asymptotic value of  $\Delta A_x$  (mixture) (i.e., if 100% of molecules were labeled at the  $p$  calculated), corrected for incomplete ion spectrum sampling, if necessary (see further discussion).

When all of the ions in the mass isotopomer spectrum are not monitored, there is no effect on calculation of  $p$ , but an effect on  $f$  may be observed, unless the calculation algorithm is corrected for the proportion of ions sampled from the labeled and unlabeled populations in the mixture of polymers (see text of article). Because of unequal molar contributions to the ion spectrum sampled, a mole of natural abundance molecules will be overrepresented in the envelope sampled relative to a mole of enriched molecules, and the assumption of linearity for mixtures (Eq. A7a) no longer applies. It is therefore necessary to weight the contributions from different populations. A simple algebraic correction can be formulated

If  $R_{(B)}$  = ratio or proportion of ions in the total envelope monitored, at baseline

$R_{(E)}$  = ratio or proportion of ions in the total envelope monitored, for enriched molecules

$M_{1(\text{meas})}$  = measured  $M_1$  fractional abundances within the spectrum sampled

$M_{1(B)}$  = True (theoretical) proportion of  $M_1$  ions in the total envelope, at baseline

$M_{1(E)}$  = True (theoretical) proportion of  $M_1$  ions in the total envelope, for enriched molecules, then:

## Equation A9b

$$f = \frac{[R_{(B)} \times M_{1(\text{meas})}] - [M_{1(B)}]}{[M_{1(E)} - M_{1(B)}] + [(R_{(B)} - R_{(E)}) \times M_{1(\text{meas})}]}$$

The derivation is as follows

$$\begin{aligned} M_{1(\text{meas})} &= \frac{[f \cdot M_{1(B)}] + [(1 - f) \times M_{1(E)}]}{f \cdot (R_E - R_B) + R_B} \\ &= \frac{f[M_{1(E)} - M_{1(B)}] + M_{1(B)}}{f(R_E - R_B) + R_B} \\ f(R_E - R_B) + R_B &= \frac{f[M_{1(E)} - M_{1(B)}]}{M_{1(\text{meas})}} + \frac{M_{1(B)}}{M_{1(\text{meas})}} \end{aligned}$$

$$f \left[ (R_E - R_B) - \left( \frac{M_{1(E)} - M_{1(B)}}{M_{1(\text{meas})}} \right) \right] = \left[ \frac{M_{1(B)}}{M_{1(\text{meas})}} - R_B \right]$$

$$\begin{aligned} f &= \frac{R_B - \frac{M_{1(B)}}{M_{1(\text{meas})}}}{\left[ \frac{M_{1(E)} - M_{1(B)}}{M_{1(\text{meas})}} \right] - (R_E - R_B)} \\ &= \frac{[R_B \cdot M_{1(\text{meas})}] - M_{1(B)}}{[M_{1(E)} - M_{1(B)}] + [(R_B - R_E) \times M_{1(\text{meas})}]} \end{aligned}$$

when  $R_{(B)} = R_{(E)} = 1.0$  (i.e., all ions are monitored, for both populations), the equation reduces to

$$f = \frac{M_{1(\text{meas})} - M_{1(B)}}{M_{1(E)} - M_{1(B)}} = \frac{\Delta A_1 (\text{mixture})}{\Delta A_1^{\infty}}$$

So the equation reduces to the standard form (Eq. A9a), as it should, when the complete ion spectrum is monitored.

Address for correspondence and reprint requests: M. K. Hellerstein, Dept. of Nutritional Sciences, Univ. of California, Berkeley, CA 94720 (E-mail: march@nature.berkeley.edu).

Received 1 September 1998; accepted in final form 19 February 1999.

## REFERENCES

1. Aarsland, A., D. Chinkes, and R. R. Wolfe. Contributions of de novo synthesis of fatty acids to total VLDL-triglyceride secretion during prolonged hyperglycemia/hyperinsulinemia in normal man. *J. Clin. Invest.* 98: 2008–2017, 1996.
2. Benyon, J. H. *Mass Spectrometry and its Application to Organic Chemistry*. New York: Elsevier, 1960.
3. Bergner, E. A., and W.-N. P. Lee. Testing gas chromatographic/mass spectrometric systems for linearity of response. *J. Mass Spectrom.* 30: 778–780, 1995.
4. Byerley, L. O., S. Bassilian, E. A. Bergner, and W. N. P. Lee. Use of D<sub>2</sub>O to quantitate cholesterol and fatty acid synthesis in tumor-bearing rats. *FASEB J.* 7: A289, 1993.
5. Chinkes, D. L., A. Aarsland, J. Rosenblatt, and R. R. Wolfe. Comparison of mass isotopomer dilution methods used to calculate VLDL production in vivo. *Am. J. Physiol.* 271 (Endocrinol. Metab. 34): E373–E383, 1996.
6. Deines, P. The isotopic composition of reduced organic carbon. In: *Handbook of Environmental Isotope Geochemistry. The Terrestrial Environment*, edited by P. Fritz and J. C. Fontes. Amsterdam: Elsevier, 1980, vol. 1, chapt. 9.
7. Dekker, E., M. K. Hellerstein, J. A. Romijn, R. A. Neese, N. Peshu, E. Endert, K. Marsh, and H. P. Sauerwein. Glucose homeostasis in children with falciparum malaria: precursor supply limits gluconeogenesis and glucose production. *J. Clin. Endocrinol. Metab.* 82: 2514–2521, 1997.
8. Des Rosiers, C., F. David, M. Garneau, and H. Brunen-graber. Nonhomogenous labeling of liver mitochondrial acetyl-CoA. *J. Biol. Chem.* 266: 1574–1578, 1991.
9. Dietschy, J. M., and M. S. Brown. Effect of alterations of the specific activity of intracellular acetyl-CoA pool on apparent rates of hepatic cholesterogenesis. *J. Lipid Res.* 15: 508–516, 1974.
10. Fagerquist, C., and J.-M. Schwarz. Gas-phase acid-base chemistry and its effects on mass isotopomer abundance measurements of biomolecular ions. *J. Mass Spectrom.* 33: 144–153, 1998.
11. Fagerquist, C. K., R. A. Neese, and M. K. Hellerstein. Molecular ion fragmentation and its effects on mass isotopomer abundances of fatty acid methyl esters ionized by electron impact. *J. Am. Soc. Mass Spectrom.* 10: 430–439, 1999.
12. Faix, D., R. A. Neese, C. Kletke, S. Walden, D. Cesar, M. Coutlangus, C. H. L. Shackleton, and M. K. Hellerstein. Quantification of periodicities in menstrual and diurnal rates of

- cholesterol and fat synthesis in humans. *J. Lipid Res.* 34: 2063–2075, 1993.
13. Hellerstein, M. K. Relationship between precursor enrichment and ratio of excess  $M_2$ /excess  $M_1$  isotopomer frequencies in a secreted polymer. *J. Biol. Chem.* 266: 10920–10924, 1991.
  14. Hellerstein, M. K., M. Christiansen, S. Kaempfer, C. Kletke, K. Wu, J. S. Reid, N. S. Hellerstein, and C. H. L. Shackleton. Measurement of de novo hepatic lipogenesis in humans using stable isotopes. *J. Clin. Invest.* 87: 1841–1852, 1991.
  15. Hellerstein, M. K., and R. Neese. Mass isotopomer distribution analysis: a technique for measuring biosynthesis and turnover of polymers. *Am. J. Physiol.* 263 (Endocrinol. Metab. 26): E988–E1001, 1992.
  16. Hellerstein, M. K., J.-M. Schwarz, and R. A. Neese. Regulation of hepatic de novo lipogenesis in humans. *Annu. Rev. Nutr.* 16: 523–557, 1996.
  17. Hellerstein, M. K., K. Wu, S. Kaempfer, C. Kletke, and C. H. L. Shackleton. Sampling the lipogenic hepatic acetyl-CoA pool in vivo in the rat. Comparison of xenobiotic probe to values predicted from isotopomeric distribution in circulating lipids and measurement of lipogenesis and acetyl-CoA dilution. *J. Biol. Chem.* 266: 10912–10919, 1991.
  18. Hentze, M. W. Determinants and regulation of cytoplasmic mRNA stability in eucaryotic cells. *Biochim. Biophys. Acta* 1090: 281–292, 1991.
  19. Hetenyi, G. Correction for the metabolic exchange of  $^{14}\text{C}$  for  $^{12}\text{C}$  atoms in the pathway of gluconeogenesis in vivo (Abstract). *Federation Proc.* 41: 109, 1982.
  20. Kelleher, J. K., A. T. Kharroubi, T. A. Aldaghlis, I. B. Shambat, K. A. Kennedy, A. L. Holleran, and T. M. Masterson. Isotopomer spectral analysis of cholesterol synthesis: applications in human hepatoma cells. *Am. J. Physiol.* 266 (Endocrinol. Metab. 29): E384–E395, 1994.
  21. Kelleher, J. K., and T. M. Masterson. Model equations for condensation biosynthesis using stable isotopes and radioisotopes. *Am. J. Physiol.* 262 (Endocrinol. Metab. 25): E118–E125, 1992.
  22. Lee, W. N. P., E. A. Bergner, and Z. K. Guo. Mass isotopomer pattern and precursor-product relationship. *Biol. Mass. Spectrom.* 21: 114–122, 1992.
  23. Mills, M., J. Chen, R. Neese, M. Hellerstein, and J.-M. Schwarz. Mass isotopomer distribution analysis (MIDA) for measurement of gluconeogenesis (GNG) under conditions of hyperglycemia in streptozotocin (stz) diabetic rats (Abstract). *Diabetes* 47, Suppl. 1: A287, 1998.
  24. Mook, W. G., and P. M. Grootes. The measuring procedure and corrections for the high-precision mass-spectrometric analysis of isotopic abundance ratios, especially referring to carbon, oxygen and nitrogen. *Int. J. Mass Spectrom. Ion Phys.* 12: 273–298, 1973.
  25. Neese, R. A., D. Faix, C. Kletke, K. Wu, A. C. Wang, C. H. L. Shackleton, and M. K. Hellerstein. Measurement of endogenous synthesis of plasma cholesterol in rats and humans using MIDA. *Am. J. Physiol.* 264 (Endocrinol. Metab. 27): E136–E147, 1993.
  26. Neese, R. A., D. Faix, J.-M. Schwarz, S. M. Turner, D. Vu, and M. K. Hellerstein. Measurement of gluconeogenesis and rate of appearance of intrahepatic triose-phosphate and its regulation by substrates by mass isotopomer distribution analysis (MIDA). Testing of assumptions and potential problems. *J. Biol. Chem.* 270: 14452–14463, 1995.
  27. Neese, R. A., and M. K. Hellerstein. Appendix. Calculations for gluconeogenesis by MIDA. *J. Biol. Chem.* 270: 14464–14466, 1995.
  28. Papageorgopoulos, C., K. Caldwell, C. H. L. Shackleton, H. Schweingruber, and M. K. Hellerstein. Measuring of protein synthesis by mass isotopomer distribution analysis (MIDA). *Anal. Biochem.* 267: 1–16, 1999.
  29. Park, O.-J., D. Cesar, D. Faix, K. Wu, C. H. L. Shackleton, and M. K. Hellerstein. Mechanisms of fructose-induced hypertriglyceridemia in the rat: activation of hepatic pyruvate dehydrogenase (PDH) through inhibition of PDH kinase. *Biochem. J.* 282: 753–757, 1992.
  30. Patterson, B. W., and R. R. Wolfe. Concentration dependence of methyl-palmitate isotope ratios by electron impact ionization gas chromatography/mass spectrometry. *Biol. Mass Spectrom.* 22: 481–486, 1993.
  31. Peroni, O., V. Large, and M. Beylot. Measuring gluconeogenesis with  $[2-^{13}\text{C}]$ glycerol and mass isotopomer distribution analysis of glucose. *Am. J. Physiol.* 269 (Endocrinol. Metab. 32): E516–E523, 1995.
  32. Previs, S. F., C. A. Fernandez, D. Yang, M. V. Soloviev, D. France, and H. Brunengraber. Limitations of the mass isotopomer distribution analysis of glucose to study gluconeogenesis substrate cycling between glycerol and triose phosphates in liver. *J. Biol. Chem.* 270: 19806–19815, 1995.
  33. Sackett, W. M. Stable carbon isotope studies on organic matter in the marine environment. In: *Handbook of Environmental Isotope Geochemistry. The Terrestrial Environment*, edited by P. Fritz and J. C. Fontes. Amsterdam: Elsevier, 1980, vol. 1, chapt. 4, p. 141–178.
  34. Schumann, W. C., I. Magnusson, V. Chandramouli, K. Kumaran, J. Wahren, and B. R. Landau. Metabolism of  $[2-^{14}\text{C}]$ acetate and its use in assessing hepatic Krebs cycle activity and gluconeogenesis. *J. Biol. Chem.* 266: 6985–6990, 1991.
  35. Seeman, J. I., H. V. Secor, R. Disselkamp, and E. R. Bernstein. Conformational analysis through selective isotopic substitution: supersonic jet spectroscopic determination of the minimum energy conformation of o-xylene. *J. Chem. Soc. Chem. Commun.* 1992: 713–714, 1992.
  36. Siler, S. Q., R. A. Neese, M. P. Christiansen, and M. K. Hellerstein. The inhibition of gluconeogenesis following alcohol in humans. *Am. J. Physiol.* 275 (Endocrinol. Metab. 38): E897–E907, 1998.
  37. Srere, P. Complexities of metabolic regulation. *Trends Biochem. Sci.* 19: 519–520, 1994.
  38. Waterlow, J. C., P. J. Garlick, and D. J. Millward (Editors). *Protein Turnover in Mammalian Tissues and in the Whole Body*. Amsterdam: North Holland, 1978.
  39. Wolfe, R. R. *Tracers in Metabolic Research. Radio-Isotope and Stable Isotope/Mass Spectrometric Methods*. New York: Liss, 1984, p. 119–144.
  40. Zilversmit, D. B. The design and analysis of isotope experiments. *Am. J. Med.* 29: 832–848, 1960.